AV COMMUNICATION REVIEW

AUDIO VISUAL

A. A. Lumsdaine — Teaching Machines and Self-Instructional Materials

Loran C. Twyford, Hildreth McAshan, Paul Messier, Philip Rowley — Behavioral and Factual Analysis

Lionel C. Barrow, Jr., Bruce H. Westley — Intelligence and the Effectiveness of Radio and Television

Malcolm S. MacLean, Jr., Hans H. Toch — The Group Interview as an Audience Reaction Measure

Departments

- Film
- Book Reviews
- Research Abstracts

SUMMER 1959

VOL. 7
NO. 3

AUDIO-VISUAL COMMUNICATION REVIEW

WILLIAM H. ALLEN, *Editor*
System Development Corporation
Santa Monica, Calif.

KENNETH NORBERG,
Departmental Editor
U. S. Office of Education

MICKEY BLOODWORTH, *Managing Editor*
Department of Audio-Visual Instruction

ANN MARIE McGOVERN, *Editorial Assistant*
NEA Publications Division

Editorial Board

Walter S. Bell
Atlanta (Ga.) Public Schools

Amo De Bernardis
Portland (Oreg.) Public Schools

James D. Finn
University of Southern California

Fred F. Harcleroad
Alameda (Calif.) State College

Charles F. Hoban
University of Pennsylvania

Anna L. Hyer
Dept. of Audio-Visual Instruction

L. C. Larson
Indiana University

Wesley C. Meierhenry
University of Nebraska

A. W. VanderMeer
Pennsylvania State University

Paul W. F. Witt
Teachers College,
Columbia University

Copyright 1959 by
Department of Audio-Visual Instruction, National Education Association

For complete Library of Congress cataloging information,
see last numbered page.

The Audio-Visual COMMUNICATION REVIEW is issued quarterly. Subscription $4 per year. $3 of every $8 DAVI membership fee is allocated for subscription to AVCR. Single copy $1.25. Entered as second class matter in Washington, D. C.

Original manuscripts, books for review, and editorial communications should be sent to the Editor: WILLIAM H. ALLEN, System Development Corporation, 2500 Colorado Avenue, Santa Monica, California. (Mark envelope "Personal.")

Subscriptions and business communications should be sent to: MICKEY BLOODWORTH, Managing Editor, Department of Audio-Visual Instruction, National Education Association, 1201 16th Street, N.W., Washington 6, D. C.

AUDIO VISUAL
COMMUNICATION REVIEW

VOLUME VII, NUMBER 3 SUMMER 1959

CONTENTS

Articles

• TEACHING MACHINES AND SELF-INSTRUCTIONAL MATERIALS 163
A. A. LUMSDAINE

• BEHAVIORAL AND FACTUAL ANALYSIS..... 182
LORAN C. TWYFORD
HILDRETH McASHAN
PAUL MESSIER
PHILIP ROWLEY

• INTELLIGENCE AND THE EFFECTIVENESS OF RADIO AND TELEVISION..... 193
LIONEL C. BARROW, JR.
BRUCE H. WESTLEY

• THE GROUP INTERVIEW AS AN AUDIENCE REACTION MEASURE 209
MALCOLM S. MacLEAN, JR.
HANS H. TOCH

Film

• FILM EVALUATION PROCEDURE USED AT INDIANA UNIVERSITY AS APPLIED TO A REPRESENTATIVE SELECTION OF RECENT EDUCATIONAL MOTION PICTURES..... 220
CAROLYN GUSS

Book Reviews

- MAY, MARK A.; LUMSDAINE, ARTHUR; AND OTHERS
  Learning from Films 228
  Reviewed by PAUL R. WENDT

Research Abstracts

TELEVISION

- *The Impact of an Educational Television Program on Its Audience—"Books and Ideas"* 237
  LAWRENCE MYERS, JR.

FILM

- *Psychiatric Patients' Audience Reactions to Types of Motion Pictures* 237
  JOHN W. WHITMYRE

PHOTOGRAPHS

- *Binocular Resolution and Perception of Race in South Africa* 238
  THOMAS F. PETTIGREW AND OTHERS

TEACHING MACHINES

- *Green Light Rater* 239
  L. C. TWYFORD AND OTHERS

TEACHING MACHINES AND SELF-INSTRUCTIONAL MATERIALS

• A. A. LUMSDAINE

This description of the current development of teaching machines presents valuable background information for understanding the part such self-teaching devices may play in modifying teaching techniques. A. A. Lumsdaine is currently the program director for training and education at the American Institute for Research, located in Pittsburgh, Pennsylvania. His most recent book, Learning from Films, *edited jointly with Mark A. May, was published by the Yale University Press in 1958.*

THIS PAPER discusses some of the developments that have occurred during the past few years in the construction of mechanical devices for individual self-instruction. It has become increasingly apparent that such devices and the instructional programs they present are likely to have very important implications for educational practice. It is even possible that in the relatively near future they could modify profoundly many of our basic conceptions of instructional method and of the role of teaching aids and audio-visual materials in the educational process. Some basic issues underlying the educational applications of automation and other technological advances have been described in several provocative papers by Skinner (21, 22), Ramo (18), and Finn (4, 5).

Teaching machines and the instructional programs used in them have developed partly on the basis of a large body of work by psychologists in conducting experimental studies of learning, and the features the machines incorporate reflect important principles of effective learning to which these studies have pointed. The present paper does not attempt to summarize this underlying experimental data, but instead concentrates on developments in devices and materials. In these developments an attempt has been made to implement principles of effective learning in ways which will be indicated in relation to features of the various devices.

All of the devices that have been called "teaching machines" represent some form of variation on what can be called the "Socratic" method of teaching. That is, they present the individual student with programs of questions and answers, problems to be solved, or exercises to be performed. In addition, however, they always provide some type of automatic feedback or correction to the student so that he is immediately informed of his progress at each step and given a basis for correcting his errors. They thus differ from films, TV, and most other audio-visual media as ordinarily utilized because of three important properties:

- First, continuous active student response is required, providing explicit practice and testing of each step of what is to be learned.
- Second, a basis is provided for informing the student with minimal delay whether each response he makes is correct, leading him directly or indirectly to correction of his errors.
- Third, the student proceeds on an individual basis at his own rate—faster students romping through an instructional sequence very rapidly, slower students being tutored as slowly as necessary, with indefinite patience to meet their special needs.

The devices thus represent a way of providing a preprogrammed study-practice combination which simulates, in partially or fully automated fashion, the functions of a private tutor in recitation and practice, with immediate correction of errors and feedback to the student.

The device shown in Figure I illustrates many of the features of teaching machines. The panel at the left of Figure I gives a close-up view of the display with which the student shown working at the right is faced. This device was designed by Besnard, Briggs, et al (2) as part of an Air Force research and development program. The device, called the "Subject-Matter Trainer," represents a further evolution of some of the simpler devices which were developed and studied earlier by Professor Sidney L. Pressey and some of his students at Ohio State University. About half a dozen of these machines have been constructed and tried out experimentally for instruction of technical specialists in the Air Force.

In essence, what this machine does is to present a predetermined sequence of questions, one at a time, in the aperture at the left of the panel. The questions can be answered by selecting any of about 20 response choices which are displayed either verbally or pictorially on the large panel. Beside each answer choice is a green

indicator light and a push button by which the student may identify and record the answer he selects as the correct choice. The device has considerable flexibility and can be programmed for use either as a testing machine or for any of several modes of operation for self-instruction in a variety of subject matters. It can be programmed, for example, so that the green light indicating the correct answer choice comes on at once the first time through, so that the student is guided to the correct choice. Later, he may be required to decide on a choice; if correct, the light beside the answer he has chosen so informs him, while if incorrect the light next to the answer that *is* correct comes on. Still later, he may be required to keep trying until he gets the correct answer. All the while, his performance is recorded and his errors totalled, so that he is continuously tested while he is learning. As he progresses, the machine can drop out the items he has mastered so that the student's time and interest are not dissipated by unnecessary repetition.

Figure I

"Subject-Matter Trainer." This is one of a class of automatic self-instructional devices or "teaching machines" particularly suitable for training in identifying of components, terms, and other "one-to-one" associations. The student is continuously tested as he learns.

Obviously, this is a rather expensive device. It was built, however, for Air Force training situations in which its cost would be very minor in relation to the total training cost for technicians who used it, not to mention the cost of errors that ineffective training could occasion in the operation of a 10 million dollar bomber or missile. In addition, this particular device was designed to have considerable flexibility and a number of modes of operation because it was also planned as a research tool for investigating various features, data on which would affect the design of simpler future devices for standard training situations.

Let us now take a look at some of the smaller and somewhat simpler devices that have been developed for use in civilian educational situations.

The device at the left in Figure II might be thought of as an ancestor of all subsequent teaching machines. It was devised by Professor S. L. Pressey (14, 15, 16) over 30 years ago. Originally, it was conceived primarily as an automatic *testing* device, but it soon became apparent that its *instructional* properties were also of great interest. In the sketch labelled Figure IIA is seen a window, at the left, in which a multiple-choice question appears. For example, in a course in elementary physics or electricity, the question might ask the student to identify the proper symbol to represent electric current or to solve a simple problem applying Ohm's Law. His basis for doing so might be the prior study of his textbook, or one of Harvey White's physics TV/film lessons, or some other source—*including what he had learned from answering preceding questions in the series.* In any case, suppose he chooses the proper answer and presses the third push button corresponding to this answer. If he does so, the machine signifies the correctness of his answer by presenting the next question or problem. But if he is wrong, the question stays in the window, an error is tallied on a counter at the back of the device, and he must try again until he selects the right answer before the next question will be presented.

Some of the other devices developed by Pressey and his students were further simplified so that the 4-choice answer apparatus was used with an ordinary mimeographed test form. This introduces the disadvantage that there is less control over the sequencing of information presented to the student—though these devices were keyed so that the student had to *answer* the questions in prescribed order, he still could look ahead or back in the series of ques-

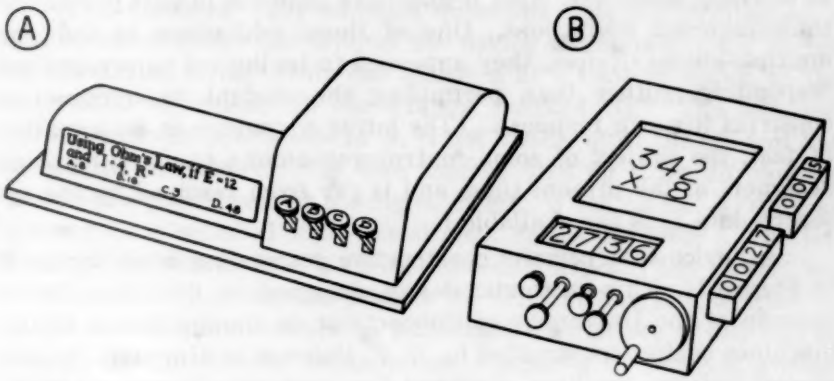


Figure II

Two simple individually paced "teaching machines" which have automatic self-scoring features as in the device shown in Figure I.

tions and thereby obtain uncontrolled or nonprogrammed clues to the answering of a given question. How important this disadvantage really is has not, curiously enough, been given much attention, and research is needed to investigate this problem.

In similar devices developed by Pressey and others, the mechanical gadgetry was simplified by the use of punchboards to replace the mechanical push-button arrangements. In the punchboard device (not shown) the student simply punched his pencil point through one of four holes in a small punchboard corresponding to the answer he selected. If right, a key-card under the punchboard holes would let the pencil penetrate to perforate the answer sheet; if wrong, the pencil point would be stopped, indicating an error to the student—who then tried another answer.

Still a different device for self-scoring and immediate feedback was devised by Peterson (11) by use of chemically treated answer sheets and a special marking pencil. These identified the correct answer by turning color when the student marked it.

For some reason these multiple-choice devices, though shown by experimental tests to be quite effective aids to learning, never caught on widely during the two decades following the initial development of Pressey's first machines. This may be partly because the times were not ripe for their acceptance, and partly that they were conceived primarily as testing devices and only secondarily

as teaching machines. Also, it may have been due in part to some of their inherent limitations. One of these limitations is that, as multiple-choice devices, they appeared to be limited to *recognition* responding, rather than permitting the student to compose or construct his own response. (The latter advantage is, as a matter of fact, the subject of some controversy among teaching-machine designers at the present time, and is far from resolved by the research data thus far available.)

A device that permits constructive responding is shown at B in Figure II. This particular device, designed by Professor David Zeaman at the University of Connecticut, is similar to one of the machines earlier constructed by B. F. Skinner at Harvard. It is a special purpose machine, designed for teaching elementary arithmetic to grade-school children.

A problem appears in the large window on the face of the machine. The problem may be of any degree of difficulty from 2 x 2 up to a problem in long division or square root. The student, after doing whatever supplementary mental or scratch-paper figuring he finds necessary, composes his answer by moving the four plungers at the front of the machine. These operate the dials to show his answer in the aperture below the problem. (In this case, 2736, the correct answer, has been set in.) When he is satisfied with his answer the student turns the crank at the right. If the answer is correct, the machine advances to the next problem by moving a tape on which the problems are printed. If wrong, it scores an error and leaves the problem unchanged so that he must try again until he gets the correct answer before he can proceed.

Several similar devices have been built, some with an extended repertoire of response-construction including letters as well as numbers, to permit their use for teaching spelling and other verbal subjects. In these devices, as in Pressey's 4-answer multiple-choice machine and Brigg's more elaborate 20-answer multiple-choice apparatus, the correctness of the answer is scored automatically.

Most devices with more extensive response repertoires have sacrificed automaticity of scoring to keep the instrumentation from becoming too complex (or even from exceeding the present state of the art). In these devices, the student himself judges whether the answer he has given is correct, and instructs the machine accordingly. Perhaps the best known device in use thus far is one devised by Professor Skinner at Harvard and shown at A in Figure III.

This machine (22) employs questions that are printed on a disc mounted under the cover of the device and shown in the sketch by the dotted circle. Only one question at a time is visible to the student as it appears in the window marked "Q." The adjacent correct answer, just above at A, is not visible until the student has completed his answer. He does this by writing it on an exposed frame of a paper tape at the right, at R1. When he has written an answer he is satisfied with, the student raises the lever at the left front of the machine. This simultaneously exposes the correct answer at A, and moves the answer the student has just written up under a transparent cover to R2, where he can see it but can no longer change it. The student now compares the correct answer with his own and decides whether he was right or wrong. If he decides he was correct, he moves the lever over to the right. This

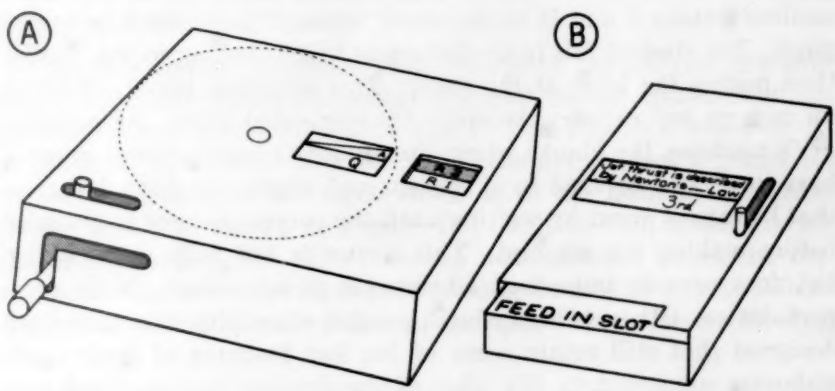


Figure III

Two "teaching machines" using written responses which are scored by the student on the basis of comparing his written answer with the correct answer.

marks his answer tape to show that he scored himself correct on that question, and also operates a mechanism that reduces the number of times that question will appear again. (The number of correct answers required for each question may be either one or two with this machine, depending on the way it is set.) Whether the answer was right or wrong, the next question appears as soon as the lever is returned to the lower position, and the student goes on with the new question.

The most elaborate aspect of the internal mechanism in this device is the part that allows for dropping out questions that have been answered once or twice correctly. One of the numerous questions on which we need research as a basis for future designs is that of how important this drop-out feature is for various kinds of subject matter and kinds of students. This question interacts with the number of questions included in any one cycle. For the machine in Figure IIIA, the cycle comprises the 30 questions on one disc. In the much simpler device shown at B in Figure III, a shorter cycle of only five questions is used, with no drop-out arrangement. In this device, built by Douglas Porter (12, 13) at Harvard, questions on mimeographed or dittoed sheets are fed into the device at the bottom and appear in the window. The student marks his answer directly on this sheet. For example, in a program used for teaching elementary spelling in a nonrote, context approach, the device might present the sentence "Rain came down in a steady pour," with the omitted letters T and D in the word "steady" to be filled in by the pupil. The student fills in the letters he thinks belong in the blanks, then pushes the lever at the right. This advances the sheet about an inch up and exposes the correctly completed word. As in Skinner's machine, the blanks where the student's response was entered have now been covered by a transparent plastic or glass cover, so that he cannot cheat by waiting until the correct answer is revealed before making his attempt. This device is not fully self-scoring, but does provide important advantages in a low-cost, fairly compact device. Other devices of still greater simplicity also have been designed that still retain some of the key features of their more elaborate counterparts (6). One of the simplest is a so-called programmed textbook devised by Professors Glaser and Homme at the University of Pittsburgh.

Glaser, Homme, and Evans (7) employ response-and-correction sequences provided by a special, nonconventional form of "programmed" question-and-answer book, in which the answer to each question is always given on the following page. These devices are especially attractive in terms of cost, but they lack some of the features of stimulus control, and the fun of learning and being quizzed by a semiautomatic "gadget" or device—which is part of the appeal that slightly more complex devices seem to have for students. Also, they do not have the self-scoring features that are capable of relieving the teacher of the drudgery of quiz scoring

and, since the answer appears on the succeeding page in each case, they are not cheat-proof even as practice devices.

How important the self-scoring and cheat-proof features are is partly an administrative question—i.e., how much do we want to pay for getting an automatically scored quiz paper—and partly a question for research—i.e., how much is learning benefited if the device is constructed so that the student cannot readily beat it? There is some evidence on the latter question, but more research will be needed before we have anything approaching a definitive answer.

The problem of making a device fully automatic and cheat-proof runs counter to the requirements for making it inexpensive and yet allowing a variety of kinds of free response to be made by the student. One experimental research device constructed by E. Z. Rothkopf (20) to experiment with a variety of response repertoires is shown in Figure IV.

In this device, extended multiple-choice capabilities are provided, together with automatically scored verbal composition responses of any of the 26 letters and 10 numbers in a word or expression of up to some 30 symbols. These are entered through the use of the sliders at the right, much as in the Zeaman numerical device that was shown in Figure IIB. In addition, a self-scoring tracer stylus is provided, which can automatically score the correctness of a student's response in tracing a route on a map, following an electrical circuit diagram, constructing a simple drawing, or the like.

Both the Skinner and Porter machines and the Glaser and Homme so-called "paper machine," or the programmed Socratic textbook, have been tried out with a number of subject matters experimentally—the former in full-year courses in some cases, the latter for shorter segments of a few experimental lessons.

As was stated earlier, all of the developments in preprogrammed self-instructional materials, and devices for presenting them, are based fundamentally on Socratic question-and-answer, or problem-and-solution, methods of teaching. Typically, they proceed in small steps of graded difficulty, so that mastery of concepts, understanding, and skills are gradually built up as the student proceeds through the program. An element that is probably crucial in their effectiveness is that of nearly continuous active student response that is monitored, guided, and corrected at each step in

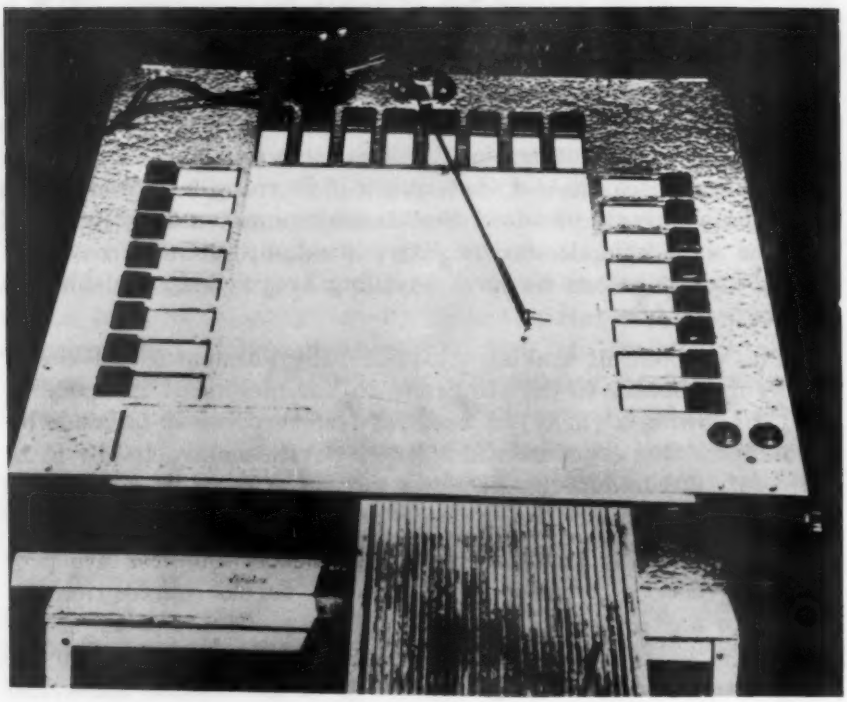


Figure IV

An experimental teaching device designed for research study of several modes of student response.

such a way as to insure mastery of one step before going on to the next.

In addition, optimum pacing to fit each student's rate of progress can be provided, and the sequencing of instructional materials and amount of repetitive practice employed can be governed by the proficiency shown at each step by the individual student. The slower learner is thus gradually brought up to a satisfactory level of accomplishment, while the quicker student, after reaching a satisfactory level of attainment as fast as his abilities permit, can then, if desired, proceed to more advanced material.

The programs of self-quizzing, self-teaching items used in practice machines and related devices are of great interest, and the design of these programs is one of the crucial factors in the successful utilization of automated individual tutoring methods. Perhaps the most important single factor in constructing programs

is the use of prompts or coaching cues, of varying degrees of subtlety and indirectness, which are used to keep the learner responding correctly and to keep up his motivation by insuring successful performance. The programming of such cues or prompts is arranged with initial maximum use to insure elicitation of the desired responses by the students, followed by so-called "fading" or "vanishing" of the prompts in later items of a sequence so that the learner is gradually freed from dependence on them.

In the instructional programs devised by Skinner, Glaser, Homme, Porter, and others, great pains are taken to develop desired response patterns by using a wide variety of contextual and semantic prompts. One simple example of the kinds of verbal prompts that can be used is partial presentation of a word, with omitted letters to be filled in—the letters present varying from nearly all present (maximal cueing) initially, to all or nearly all absent (minimal cueing) terminally.

For programs that are designed to build up mastery of complex concepts rather than to develop simpler skills, other techniques are useful. Two examples are presented in Table 1.

In the first example, analogy is used to lead the student to the appropriate response in context for the first time. Almost inevitably, he will respond with the desired term to complete the statement that a "rise in filament temperature will *increase* electron emission from the filament. . . ."

In the second example, pre-existing word habits are used in a different form of cueing: The structure and meaning of the sentence make it very likely that the student will respond correctly by saying that "the *first* component we encounter is the RF detector."

TABLE 1.—SOME PROMPTING TECHNIQUES USED WITHIN FRAMES OF TEACHING MACHINE PROGRAMS

I. *Example of Prompting by Analogies or Examples:*
Just as a rise of temperature increases gas pressure, a rise in filament temperature will.....?.....electron emission from the filament of a vacuum tube.

II. *Example of Prompting by Word Habits:*
When we start to trace a signal from the antenna to the viewing tube, the.....?..... component we encounter is the RF detector.

Another general method of providing prompts which foster correct responding in context is through dependence on stimuli provided in immediately preceding frames of the sequence.

Table 2 shows a brief 4-item excerpt from a program in electrical fundamentals. The items would normally be seen only one at a time. Assume the student responds correctly to Item 62 with the word *"transformer,"* on the basis of what he has learned in preceding sequences. If he does not, the correct answer will, of course, be provided after he has tried to answer the item. Cues for the response *"coils"* for Item 63 are provided by the statement just previously read in Item 62; similarly, the terms *"primary"* and *"secondary"* required in Item 64 will have just been used in Item 63, and so on. A number of verbal prompting techniques have been described by Smith (23).

TABLE 2.—PROMPTING BY CUES FROM PRECEDING ITEMS IN A TEACHING MACHINE PROGRAM

62. A.....?....., used to control voltages, consists of two coils of wire, wound on a core.
63. Input and output voltages of a transformer are governed by the number of turns of wire in the primary and secondary.....?.....?
64. The ratio of the number of turns in the .....?..... and in the .....?.....coils determines the ratio of input and output voltages.

The techniques of providing prompting cues may be subtle or more obvious, and can be used for eliciting responses to nonverbal stimuli as well as for purely verbal patterns. An instructive example is cited by Skinner in his 1958 article in *Science:*

> When a student describes the geography of part of the world or the anatomy of part of the body, or names of plants and animals from specimens or pictures, verbal responses are controlled by nonverbal stimuli. In setting up such behavior the student is first asked to report features of a fully labeled map, picture, or object, and the labels are then vanished. In teaching a map, for example, the machine asks the student to describe spatial relationships among cities, countries, rivers, and so on, as shown on a fully labeled map. He is then asked to do the same with a map in which the names are incomplete, or possibly lacking. Eventually he is asked to report the same relationships with no map at all. If the material has been well programmed, he can do so correctly.

Reviewing the foregoing examples, it is clear that a great deal of ingenuity and skill may be expended in constructing good pro-

grams, and also that the design of both devices and programs suggests a number of problems on which research is needed to provide firmer guidelines for future development. Some of the more important problems are the question of the size of steps in the programs, the optimal use of prompts, and the rate at which the prompts are to be eliminated or "vanished."

The principles used in most of the programs constructed by Skinner, Gilbert, Homme and Glaser, Porter, and others appear very reasonable, but have not yet been subjected to clear-cut experimental tests. Some highly suggestive evidence is available from studies of active student response in learning from films. For example, an Air Force study by Kimble and Wulff (9) found that increasing the amount of coaching and guidance furnished to students in learning to use a slide rule resulted in superior learning. Other studies summarized by Allen (1) document the value of frequent guided response by the student. In experimenting with programmed materials for teaching several subject matters, Glaser and Homme at the University of Pittsburgh have found that, up to a point, the smaller the steps in the program, the better. How to determine that point is at present a highly empirical matter, and we need to develop a theory and procedures for predicting the best size of step to use with any subject matter.

Some very suggestive evidence related to the vanishing technique, in the case of procedural learning, was obtained in an Air Force-sponsored program of research through the theoretical and experimental work of Maccoby and Sheffield (10), who concluded that transition from initially short demonstration-and-practice segments to larger segments later on would produce better terminal performance in a nonprompted test situation than consistent use of either long or short segments throughout. (In this case the larger segments provide more remote, thus less direct, cueing for the steps of the procedure.)

The effective use of short and variable length sequences for individual instruction by filmed demonstration requires projection equipment with characteristics that permit its use as a good teaching machine and that are not at present provided by commercially available projectors. One suitable device that exists at present only in a single prototype model is shown in Figure V.

This device (8) shares with the teaching machines discussed previously the characteristics of presenting materials in a fixed

sequence, at a pace that can either be preprogrammed or determined by the responses of the individual learner using it. In this case, however, the materials presented are sound motion pictures, making the device particularly adaptable to the automated presentation of demonstrational sequences for both conceptual and procedural training. The device shown here differs from other sound motion pictures in the nature of its preprogrammable start-and-stop characteristics, and also in the fact that because its endless film loop is contained in a prethreaded magazine, it can present any desired set of film sequences from a sizable library of magazines at virtually a moment's notice, with no need for threading or handling the film by the user. Thus, the mechanics of projection are removed from the attention of the instructor or teaching assistant, being turned over to the individual student and to the machine itself.

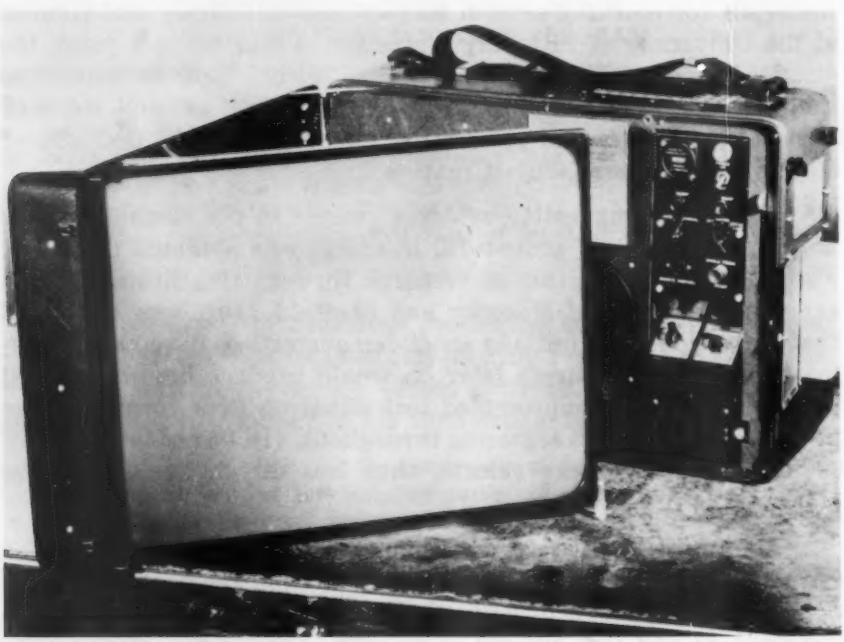


Figure V

Magazine-loaded "automated" projector. This prototype device was designed by the Air Force to provide "automated" presentation of demonstration film sequences, with flexible access to any desired film through simple insertion of prethreaded magazines.

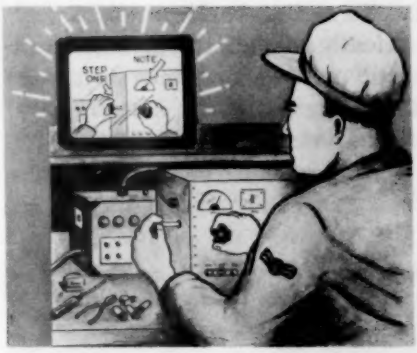


Figure VI

Use of special step-by-step film projection as a procedural training aid. A technician watching a step-by-step procedure presented in short sequences is able to copy lengthy and exacting procedures accurately with minimal training.

The sketch in Figure VI depicts the utilization of such a machine as an automated demonstrator for teaching procedural skills in the laboratory or in apprentice training of technicians on the job. The demonstrational segments to be performed by the student can be tailored in length to his ability and prior experience, so that he can perform each step and check his work as he proceeds against a picture that remains on the screen until he is ready to go on with the next step. Thus, the trainee can learn while he works and may perform useful work while he learns.

The general availability of such techniques of instruction in public schools must obviously await the development of an automatic projector, probably one with prismatic optics, that can be mass-produced at a very low unit cost. Such a projector has been for some years, and still is, one of the tools most needed for effective use of films in practical instruction.

One of the unsolved problems in the design of all forms of teaching machines is the question of program variation. In the Skinner and Briggs machines the program is fixed, except for dropout of mastered items, for some 30 frames or items, at the end of which the student may, depending on his performance and needs, go on to the next disc or go back to an earlier set of material to remedy some deficiencies in his prerequisite knowledge or skip ahead to more advanced material.

With the magazine projector shown in Figure V, the material may similarly be varied in units of five minutes or less, simply by selecting and inserting a new magazine. In some instances, even greater flexibility may be desired—so that the next problem or question presented to the student is governed by his response to each preceding question. This is likely to be particularly desirable in teaching problem-solving or diagnostic skills, and may be useful in other teaching as well. At least three methods have been used

thus far to achieve this kind of flexibility. One is the use of a digital computer (see 19) which automatically adjusts the problem difficulty to each response that the student makes. This may seem like a fantastic degree of instrumentation, but it actually has practical possibilities for future development. These, however, are outside the scope of the present paper. A second method uses a microfilm projection equipment, such as is shown in Figure VII.

Figure VII

A variable-sequence self-instructional device. The program of materials presented on the screen can be arranged in such a way that the presentation of new items of information depends upon the student's own performance.

This device (3) was developed for teaching electronic troubleshooting to Air Force technicians. It provides for a variable program in which not only the pacing but also the sequencing of instructional items can be varied, almost without limit, to adapt to qualitative differences in the requirements of individual students as indicated by their responses to previous items. This machine, and some similar ones developed experimentally in the former Air Force laboratory in Denver under the direction of R. M. Gagne, can select any sequence of questions or problems from several

hundred items stored on microfilm, projecting each item in an order that is determined by the response of the student to the preceding item or items. This particular device was designed as a tool for research in methods of teaching concepts and discriminations which Air Force technicians need to master for trouble-shooting complex electronic equipment systems. However, its basic principles are adaptable to the teaching of a wide variety of subjects, particularly those that require the acquisition of diagnostic or problem-solving skills.

A highly simplified though less flexible and less automatic variant of this device has been developed by Crowder (3) in the form of a so-called "scrambled book," in which the pages are in random order. With this "paper teaching machine," the page that the student is told to turn to next is contingent on the answer he gives to the problem or question presented on the current page.

The field of teaching machines and related devices has recently been expanding so rapidly that published information lags far behind the work accomplished and in progress. Major products of the extensive work of Skinner and his co-workers up to last year, as well as part of the antecedent work of Pressey and his students, are described in Skinner's recent article (22).

Other contributing lines of thought, experimental study, and development of programs and devices in other laboratories have been described in two recent symposia, one held at the American Psychological Association Convention in August 1958, and one held at the University of Pennsylvania in December 1958, under sponsorship of Air Force Office of Scientific Research. As yet unpublished summaries by the program chairmen of both symposia (A. A. Lumsdaine and E. H. Galanter, respectively) reflected the emphasis placed by a number of the individual participants on two things: *first*, the unresolved issues for research that clearly emerge from developments thus far, and *second*, the promise that automated learning techniques and devices afford for effecting improvements of great practical importance in education, even at the present state of the art.

REFERENCES

1. ALLEN, W. H. "Research on Film Use: Student Participation." *Audio-Visual Communication Review* 5:423-50; Spring 1957.
2. BESNARD, G. G.; BRIGGS, L. J.; MURSCH, G. A.; and WALKER, E. S. *Development of the Subject-Matter Trainer.* Air Force Personnel and Training Research Center, *Research Report* AFPTRC-TN-57-77, June 1957.

3. CROWDER, N. A. "Automatic Tutoring by Means of Intrinsic Programming." Paper presented at the Air Force Office of Scientific Research Symposium on Teaching Machines, University of Pennsylvania, December 8-9, 1958.

4. FINN, J. D. "Automation and Education: I. General Aspects." *Audio-Visual Communication Review* 5:343-60; Winter 1957.

5. FINN, J. D. "Automation and Education: II. Automatizing the Classroom." *Audio-Visual Communication Review* 5:451-67; Spring 1957.

6. GILBERT, T. F. "Some Issues Involved in Selecting an Automatic Teaching Device." A report to Bell Telephone Laboratories, Murray Hill, N. J., September 1958.

7. GLASER, R.; HOMME, L.; and EVANS, J. "An Evaluation of Textbooks in Terms of Learning Principles." Paper presented at the meetings of the American Educational Research Association, February 1959.

8. HOEHN, A. J., and LUMSDAINE, A. A. *Design and Use of Job Aids for Communicating Technical Information.* Air Force Personnel and Training Research Center, *Technical Report* AFPTRC-TR-58-7, ASTIA Doc. No. AD 152109, January 1958.

9. KIMBLE, G. A., and WULFF, J. J. *The Effect of "Response Guidance" on the Value of Audience Participation in Training Film Instruction.* HFORL Report No. 34. Washington, D. C.: USAF Human Factors Operations Research Laboratories, 1953.

10. MACCOBY, N., and SHEFFIELD, F. D. "Theory and Experimental Research on the Teaching of Complex Sequential Procedures by Alternate Demonstration and Practice." *Symposium on Air Force Human Engineering, Personnel, and Training Research.* (Edited by G. Finch and F. Cameron.) Washington, D. C.: National Academy of Sciences-National Research Council, 1958.

11. PETERSON, H. J., and PETERSON, J. C. "A New Device for Teaching, Testing, and Research in Learning." *Trans. Kansas Acad. Sci.* 33:41-47; 1930.

12. PORTER, D. A. "A Critical Review of a Portion of the Literature on Teaching Devices." *Harvard Educational Review* 27:126-47; Spring 1957.

13. PORTER, D. A. "Some Effects of Year-Long Teaching Machine Instruction in an Elementary School." Paper presented at the U. S. Air Force Office of Scientific Research Symposium on Teaching Machines, University of Pennsylvania, December 8-9, 1958.

14. PRESSEY, S. L. "Development and Appraisal of Devices Providing Immediate Automatic Scoring of Objective Tests and Concomitant Self-Instruction." *Journal of Psychology* 29: 417-47; April 1950.

15. PRESSEY, S. L. "A Machine for Automatic Teaching of Drill Material." *School and Society,* 29 (No. 645), 1927.

16. PRESSEY, S. L. "A Simple Device Which Gives Tests and Scores—and Teaches." *School and Society* 23:373-76; 1926.

17. PRESSEY, S. L. "Some History Regarding Self-Instructional Devices." Paper presented at the American Psychological Association Symposium on Automation of Teaching, Washington, D. C., August 1958.

18. RAMO, S. "A New Technique of Education." *Engineering and Science Monthly,* October 1957.

19. RATH, G. J., and ANDERSON, NANCY S. "The IBM Research Center Teaching Machine Project: (1) The Teaching of Binary Arithmetic; and (2) The Simulation of a Binary Arithmetic Teaching Machine on the IBM

650." Paper presented at the U. S. Air Force Office of Scientific Research Symposium on Teaching Machines, University of Pennsylvania, December 8-9, 1958.

20. ROTHKOPF, E. Z. "Research Problems in the Design of Materials and Devices for Automated Learning." Paper presented at the American Psychological Association Symposium on Automation of Teaching, Washington, D. C., August 1958.
21. SKINNER, B. F. "The Science of Learning and the Art of Teaching." *Harvard Educational Review* 24:86-97; 1954.
22. SKINNER, B. F. "Teaching Machines." *Science* 128:969-77; 1958.
23. SMITH, D. E. P. "Principles of Programming." Paper presented at the U. S. Air Force Office of Scientific Research Symposium on Teaching Machines, University of Pennsylvania, December 8-9, 1958.

BEHAVIORAL AND FACTUAL ANALYSIS

- LORAN C. TWYFORD
- HILDRETH McASHAN
- PAUL MESSIER
- PHILIP ROWLEY

This paper reports a new method for evaluating the effect of an audio-visual message upon the behavior of the viewers. It contrasts with the usual procedure of measuring the amount of learning or the change in attitudes. This research was carried out under the direction of Loran C. Twyford of the Highway Traffic Safety Center and Audio-Visual Center of Michigan State University under a grant from the Public Health Service (RG-5786). Dr. Twyford is now consultant on educational communications in the New York State Education Department. Hildreth McAshan, Paul Messier, and Philip Rowley are special research assistants assigned to the project.

RECENT TRENDS in education stress that learning has not really occurred until it affects our behavior. Traditionally we teach factual information and attempt to shape attitudes while anticipating that these will change behavior. Both of these methods are known to have limitations. The *Encyclopedia of Educational Research* indicates that "knowledge and behavior do not seem to be closely related" (1). There is also considerable question as to the effect of attitudes upon behavior. The purpose of this research was to determine the various factors that result in changes in behavior, and to find out which of these factors can be modified to result in improved behavior. This research deals with problems of traffic safety and has as one of its objectives the reduction in the number of casualties and serious accidents on our highways. A number of methods have been used for analyzing the content of films and other messages. A well-known method is content analysis, in which a record is made of the number of references made to various categories being studied. These categories may be in terms of morality, personality, patriotism, et cetera. Direct and indirect references are recorded. Some of these may be implied through symbols and actions. This method has been used during the period

of production of a number of films to study the major themes communicated. It has also been used to analyze the character impact of actors. This type of analysis is most useful for obtaining an over-all evaluation, and it is assumed that if material is presented often enough in a positive manner that it will be more effective. A second method for analyzing a part of a message employs rating devices with which a profile is produced. This profile may indicate degrees of like-dislike, learning, interest, and other types of ratings (2).

One of the methods for component analysis developed under this research project was called *factual analysis*. This was later developed into a more complex method called *behavior analysis*, in which the gap between knowledge and behavior was reduced. The procedure involved in factual analysis is the listing of every fact presented in a film and the analysis of the facts to estimate how much learning one would expect for various target audiences.

One of the discoveries in factual analysis was that a large number of facts can be classified as background information which has been previously known by most audiences. It is not expected that these facts would affect the audience. Examples are given in Figure I. This is one of the better films analyzed. There is another category of facts which includes those not sufficiently important to be analyzed since their impact is marginal. They are called "unimportant facts." These two categories of facts may represent as much as 80 percent of the total content of the film. Figure II indicates the number of facts in these categories for several films. A 10-minute film may include from 50 to 200 facts. The category of facts which is of most interest is that of learnable facts. If facts are already known, as so many will be, there is little chance for additional learning to occur. The main purpose of factual analysis is to separate those important aspects of a total message so that a more realistic appraisal can be made of the content of the film. The term *factual analysis* has been used because it is assumed that every statement, bit of information, or generalized concept can be stated as a message or fact which can be analyzed.

This analysis of factual content led to the understanding that many facts cannot logically be assumed to induce any change in behavior. Therefore, it became necessary to study the more complex factors influencing behavior. The methodology is incorporated in what is called *behavioral analysis*.

Figure I

Behavior and Factual Analysis of the film

PRACTICE MAKES PERFECT

(10 Min.)

	BEHAVIOR CHANGE COMPONENTS												
	KNOWN FACT	ESTIMATED LEARNING	PRESENT PERFORMANCE	CHANGE OF BEHAVIOR POSSIBLE	PERSONAL JUDGMENT	NEGLIGENCE	LACK OF KNOWLEDGE	CHANCE SITUATION	OTHER PEOPLE'S BEHAVIOR	OTHER FACTORS	ESTIMATED CHANGE	PERMANENT BEHAVIOR CHANGE	IMPORTANCE
UNIMPORTANT FACTS													
1. Practice makes perfect.	0	0	0	0	0	0	0	0	0	0	0	0	0
2. It's smart because it's safe.	0	0	0	0	0	0	0	0	0	0	0	0	0
3. The experts begin by learning the ropes.	0	0	0	0	0	0	0	0	0	0	0	0	0
(18 additional similar items were obtained)													
BACKGROUND FACTS													
1. A basketball game was in progress.	0	0	0	0	0	0	0	0	0	0	0	0	0
2. The bowler made a strike.	0	0	0	0	0	0	0	0	0	0	0	0	0
3. The golf club drove the ball.	0	0	0	0	0	0	0	0	0	0	0	0	0
(11 additional similar items were obtained)													
LEARNABLE FACTS													
1. You've got to practice, not your mistakes, but the right way.	95	3	0	0	0	0	0	0	0	0	0	0	0
2. Driving a golf ball is no different than driving a car.	12	80	0	0	0	0	0	0	0	0	0	0	0
3. You can learn to drive by practice.	92	5	0	0	0	0	0	0	0	0	0	0	0
4. Bouncing wheels don't have much traction.	85	13	0	0	0	0	0	0	0	0	0	0	0
(44 additional similar items were obtained)													

BEHAVIOR FACTS

1. Don't use your brakes on a skid.	30	15	55	45	15	3	8	2	0	17 Fear		90
F. of L.: If you use your brakes on a skid, you may cause your car to skid more. Many people use their brakes as a reaction to skidding and before they realize that they have done so. Some people are afraid of skidding and will use their brakes to try and stop this.					5		1			9	15	2
2. Turn your front wheels into a skid.	50	30	40	60	0	5	50	5	0	0		50
F. of L.: Turning the front wheels into the skid will make the car come out of the skid. Many people perform this action without thinking and do not actually know that this is the proper way of stopping a skid. This is the only sure way of stopping a skid.							40				40	3
3. Try your brakes after you get out of water.	30	60	10	90	20	20	25	0	0	Lack of		20
F. of L.: This action will show whether your brakes are dry enough to stop your car safely. Trying your brakes a few times will usually dry them out if they are wet. Many people do not perform this behavior often enough to have the proper knowledge of what performance to make.							20			Recall	30	.5
4. Blowouts will happen even with new tires.	70	20	60	40	15	15	8	0	0	Disagree		60
Conv.: There is danger in a blowout, so drive slowly.					5	8	2				15	1.5
F. of L.: Most people never have a blowout. Tires now are almost blowout-proof. Blowouts may happen at slow speeds also, so why drive slowly.												
5. If a front wheel blows out, the car will skid to the direc-	65	20	40	60	35	20	5	0	0	0		38
tion of the flat tire.					10	5	1				16	.8
Conv.: If a blowout happens, it is completely out of the hands of the driver.												
F. of L.: People should always be aware of the possibility of a blowout. If a blowout occurs, try to get your car under control and then stop.												
6. Shift to a lower gear going downhill.	50	20	50	50	15	0	15	0	5	Lack of		45
(43 additional similar items were obtained)							10			Recall	20	.1
7. Slow down for water on the highway.	60	20	50	50	20	5	10	5	5	Disagree		40
(43 additional similar items were obtained)							5		3		10	.5

BEHAVIORAL ANALYSIS

When we consider all behavioral components involved in good traffic behavior, we find that the number of components is too great to be tackled in any definitive manner. For that reason, a behavioral analysis is restricted to a few behavioral components which might be presented in a single film. Once a film has been selected that appears to present desirable behavior to be acquired, it is then necessary to list every film component. This can be done easily by projecting the film and making a record of the commentary and visual events as individual statements of fact. It is necessary to convert compound sentences into the simple sentences making up the complex idea. Since this stage is complete only when every item that the film can be expected to teach or might possibly teach is included in the listing, it would appear to be a very difficult task at first glance. However, it can be accomplished in two or three hours. Certain components which are listed may be stated in a form which makes it difficult to interpret the meaning. It is necessary to make a slight conversion, such as removing the names of persons and filling in references to previous parts of the film. It is very important to convert the facts to basic facts without alteration in meaning or general intent. This may require a rewording of the fact. Where visual information has been presented in such a way as to produce an impact on the individual, this needs to be stated as a fact also.

BEHAVIORAL - FACTUAL ANALYSIS

	PRACTICE MAKES PERFECT	MOTOR MANIA	FATAL SECONDS	AND THEN THERE WERE FOUR	NIGHT DRIVING	%
UNIMPORTANT	21	47	33	10	0	19
BACKGROUND	14	54	19	90	8	31
LEARNABLE	48	19	12	73	34	31
BEHAVIORAL	48	8	8	19	34	19

Figure II

Behavioral-Factual Analysis

When all facts have been listed, they can be analyzed. As an aid to doing this, the facts should be categorized. The four categories of facts shown in Figure I are *background, unimportant, learnable,* and *behavioral.* The first two are incidental to the main message of the film and can be listed and then ignored. Background facts are those which give the setting or general conditions of the film. Unimportant facts are those peculiar to the situation depicted and not intended to be learned as such. It may be difficult to separate these categories, and it is not necessary to do so.

The remaining facts include that information which the film purports to transmit to the viewer. Of these, the most important category can be classified as *behavioral.* That is, it can be assumed that when these facts are viewed, one can expect some change in the behavior of the viewer. Those facts which appear to be important components of the film but which would not be expected to change behavior have been called *learnable facts.* Actually all facts are learnable, but these are peculiarly learnable in the mental sense: they become part of the mentality but are not actionable. They might be labeled "mental facts," "rational facts," or "facts that make up attitudes." It is these two categories of behavioral and learnable facts that must be fully analyzed.

After the facts in the film have been categorized, it becomes necessary to analyze the reaction of the viewer. How a message is received depends on the previous experience of the viewer and on the facts which he brings to bear on the message. Therefore, for each fact which may result in a change of behavior, we must list that information which we expect the learner to associate with the fact. This has been called the listing of the "facts of life," for lack of a better title. A typical example of the operation of these components is in teaching one to respect speed laws. A film might state that one should not exceed the speed limit, but the viewer of the film is forced in everyday driving habits to exceed speed limits on certain roadways. He also knows that he can exceed the limit by a few miles per hour and still not be ticketed. Therefore, based on the facts stated in the film and his knowledge of common behavior, the learner may fail to learn anything new and will certainly not change his behavior. Consequently, each fact must be treated in context with the facts of life in order to obtain an accurate estimate of anticipated behavior. The behavioral items in Figure I show how this is done for each behavioral fact. It may be desirable to do this also for learnable facts.

At this point, it is possible to begin making estimates of the impact of the film. To do this realistically, one must select the target audience for which the estimates are to be made. The necessity for selecting the proper audience is important since new learners, having little experience, are prone to accept the information as correct and final, whereas older learners bring to bear a greater amount of experience. An analysis was made of viewers of traffic safety films distributed by the Traffic Safety Film collection of the Michigan State University Highway Traffic Safety Center and the Audio-Visual Center. It was found that most films were used by driver-education groups in the high schools, and the analysis has been made for them.

One of the first estimates is that of the percentage of the viewers who know the facts. If they know the information, no learning can occur and certainly no behavior change is anticipated. This estimate can be checked by means of tests if it is not felt that the estimate is correct.

The remaining percentage consists of the learning potential. If 80 percent know a fact, then, theoretically, 20 percent could learn the information. (Figure I) This is the estimate of possible learning. If material being presented is very difficult, then the estimate of actual learning is probably somewhat less than that remaining. Although learning is probably not related directly to behavior, we can assume that there will be no behavioral changes without learning. Therefore, the estimates of learning are of value in establishing a ceiling for behavioral changes.

BEHAVIORAL ESTIMATES

To determine the possibility for a change in behavior, it is first necessary to estimate to what extent each behavior is performed. If the film fact states that one should stop completely at a stop sign, and if the average person does so 70 percent of the time, this 70 percent represents present performance—leaving a potential change in behavior of 30 percent. It is important to base these estimates on the number of incidents of the fact or performance, rather than on the percentage of time involved. From the present performance, we come to the most important estimate—the estimate of possible change in behavior. If a film does not have many components with large possibilities for change in behavior, then it cannot effectively change behavior to any extent. Now we will analyze this potential behavioral change into its components. The method

for doing this is one of working back from the reasons stated by drivers for failure to perform correctly. The most basic reasons were obtained in personal interviews of one hour's duration. Actually, these reasons may overlap and, in so far as possible, one must break down complex categories into discrete categories that may be treated individually. Some of the categories that have been obtained for driver-safety films are personal judgments, other drivers' behavior, negligence, lack of knowledge, chance situations, and other factors. Other categories would need to be developed for other types of behavior. When these categories have been set, it is possible to distribute the percentage of total possible change in behavior to the various components. It has been found that there is relatively consistent agreement in assigning these percentages when it is known what is being considered. For instance, out of the possible change of 30 percent, the reason for not conforming to recommended behavior in 10 percent of the instances may be due to personal judgment.

Although we now have estimates of the reasons why behavior has not conformed to recommendations, it is necessary to give a rating of importance for knowing the various items of information in the film. Although it is difficult to make this rating of the relative importance of various facts, it is nonetheless important that all the facts not be given equal emphasis in the final assignment.

BEHAVIOR CHANGE POSSIBILITIES

Now that all the components in the film which could affect behavior have been listed and estimates made of the reasons why behavior has not changed, it is possible to analyze the films more completely. One of the first steps is to rank the facts of the film in order of decreasing possibility for *learning* change. This will indicate which items will change the viewers mentally, even though no change in behavior may result. From such an analysis, it has been found that in a 10-minute film there may be only a few facts which affect learning. In particular, it has been found that dramatic or attitudinal type films present so much situational information that there is little time to present factual information which might change behavior. There is probably a need to avoid presenting too much factual information or more than can be adequately absorbed by the viewer. From this analysis of behavior change pos-

sibilities, it is then desirable to observe actual behavior to determine conformance with the estimates.

From this analysis, one sees that lack of knowledge is an insignificant reason for not behaving as desired. In cases we have studied, personal judgment and disagreement with laws and other drivers' behavior seem to be more important factors. To determine the relative impact of these factors, one should add together the estimates of possible change of each factor and determine which could be most easily changed to result in improved behavior.

After one has estimated values for each of the behavior change components, it is then necessary to estimate how much change in behavior one can expect as a result of seeing the film. For instance, if personal judgment accounts for 10 percent of the lack of correct behavior, one might estimate that the film will affect the viewer and produce an estimated change of 2 percent in the desired direction. Similar estimates can be made for other components. When these are totaled, one obtains an estimate of total change in behavior resulting from the presentation of that one fact which might be measured soon after seeing the film. An estimate can also be made for any permanent change that could be expected.

Here again, as in the factual analysis, one must make some rating of the importance in performing the desired behavior. Each one of the behavior change components must be analyzed to determine how behavior might change. There are a number of possibilities. If the major reason for lack of correct performance is personal judgment or disagreement with the law, then efforts should be made to give understanding to the law or change the law. If a major lack of compliance is due to chance situations beyond any control of the individual, one would not expect to obtain any improvement. If the major factor is lack of knowledge, then improved methods for presenting information might be employed. One should study the extent to which personal judgment is involved and should analyze the factors that make it up. If there is reason to analyze the effect of the film on other age levels or levels of experience, then new estimates should be made.

Estimates can be made by one person and a great deal of clarification thus obtained. If possible, it would be more desirable for a number of persons to make estimates, and the compared results of these estimates would be an index to the degree to which one should rely upon them.

LEARNING POSSIBILITIES

There are probably many instances in which learning rather than a change in behavior is assumed to be the end goal. Certain psychologists have related learning to behavior and consider them synonymous. However, the method described above can be used to estimate how much could be learned from any presentation. One can rank items on the basis of knowing the facts which will result in the reinforcement of the information presented. Another possibility of interpretation from learned material is to analyze how learning affects each of the behavior change components. Most measures of performance involve some discussion of the effects of attitudes. It is difficult to make a definition of attitudes that will satisfy educators, psychologists, and sociologists. Efforts were made to include attitudes in the above discussions, but whenever listed as separate components, it became almost impossible to obtain agreement. Therefore, when attitudes were involved, they have been subsumed in the behavioral change categories. In further experimentation, it is hoped to verify this assumption.

VALIDATION OF BEHAVIOR EFFECTS

The method described is one developed from information obtained in depth interviews—based on the assumption that people can make estimates of the various factors involved in their conformity to a particular pattern. Research has been done to validate these estimates. One method of validation was to make observations of traffic behavior. Called street-corner observations, one of these involved measurement of the compliance to the regulation of complete stops at stop signs. It was found that 93 percent of the drivers did stop completely at the signs. This bit of information therefore acts as a check on the estimates which were made of the fact in the film. Another method of observation might be for passengers to report on the behavior of the driver. A questionnaire which is also being developed to help the driver report on his own driving behavior needs to be validated.

Depth interviews were found to be valuable for assigning values to the reasons for noncompliance. Other values in the behavioral analysis were checked by tests to determine the level of knowledge and observation in order to determine behavior before and after seeing the film. The various behavior components were checked where possible. In some cases where estimates were far too

generous and it was necessary to re-evaluate the estimates, it became apparent that the film would have very little effect on behavior. With experience in estimating behavior, it was felt that the estimates would closely approach actual conditions.

SUMMARY

As a result of analyzing five films, the general conclusion was that, based on estimates made, very little behavior change could be attributed to these films, and this change must of necessity be restricted to specific bits of information as stated in the film. Further validation should verify this conclusion. The main features in any lack of behavioral change seem to be that personal judgment and other people's behavior largely control changes in behavior. What is wrong with films? Most films must be sold to nationwide audiences and must conform to laws and general standardized norms of behavior. Since actual behavior deviates so radically from recommended behavior, most films are actually giving information that is not closely related to desired behavior. What are the best ways of changing behavior? This is an important problem in traffic safety where there is a big discrepancy in present and recommended behavior. Although one may insist on safe driving practices, doing so is little help unless one is told exactly what to do. We have found a large difference in the effect of films on groups with varying amounts of experience. New learners learn a considerable amount from films and immediately translate this into correct behavior; however, behavior is soon changed as the new drivers encounter common practices on the road. Experienced drivers occasionally learn facts from films, but these are so isolated that the over-all effect is negligible.

Many research studies have shown that attitudinal type films seem to have little or no effect on behavior (3). This study has analyzed behavior into its components and has given reasons why we do not behave as we are told.

REFERENCES

1. MONROE, WALTER S., editor. *Encyclopedia of Educational Research.* Revised edition. New York: Macmillan, 1950. p. 142.
2. TWYFORD, L. C. "Profile Techniques for Program Analysis." *Audio-Visual Communication Review* 2:243-62; Fall 1954.
3. U. S. NAVAL TRAINING DEVICES CENTER. *Instructional Film Research Reports.* NAVEXOS P-1543, Vol. 2. Port Washington, L. I., N. Y.: the Center, June 1956.

INTELLIGENCE AND THE EFFECTIVENESS OF RADIO AND TELEVISION

- **LIONEL C. BARROW, JR.**
- **BRUCE H. WESTLEY**

This study presents some important insights into the relationship of intelligence to the teaching effectiveness of radio and television. A related study was published in the Winter 1959 issue of AVCR. Lionel C. Barrow, Jr., is study director in the University of Wisconsin Television Laboratory. Bruce H. Westley is associate professor of journalism, University of Wisconsin.

NUMEROUS studies have been reported on the relationship between levels of ability and the extent to which learning takes place from the audio-visual media, including television. Although these have been sometimes contradictory and often inconclusive, there is no reason to question the judgment of Hoban and Van Ormer (3) who, in a review of 12 studies made prior to 1951, state:

> It appears that . . . there is a fairly high positive correlation between intelligence and learning from . . . a pictorial-verbal medium of communication.
>
> Many . . . investigations suggest that . . . the more intelligent persons usually know more . . . when they start, make a greater absolute gain, and maintain their superiority over lower and average intelligency levels at the end.

However, the idea that the audio-visual media are particularly advantageous to the low-ability learner has appeared to gain wide currency. This may be due in part to a point made by the Hoban and Van Ormer summary, which states:

> In some cases, but certainly not universally, teaching with a film or with the aid of a film seems to bring about a greater increment in learning among those of lower intelligence than it does among those of higher intelligence. However this greater increment does not increase their total learning to the extent that it surpasses that of the average or superior groups.

Perhaps the strongest evidence that may be cited in favor of this hypothesis is that of Kanner, Runyon, and Desiderato (4), who measured the achievement of Army recruits after they had been taught basic training courses by television and by conventional instruction. High- and low-aptitude trainees were distinguished on the basis of Area I (intelligence) scores on the Army Classification Test. The result was that TV instruction produced significantly higher scores than regular instruction among low-aptitude trainees on 10 achievement tests out of 17 (seven showed no significant differences), but among high-aptitude trainees there were no significant differences. However, the authors point out that this could be largely artifactual: the high-aptitude groups tended to score so near the ceilings of the tests that there was little room for differences.

Since that time several studies have appeared which at least question the acceptability of this hypothesis. For example, two studies comparing the effectiveness of closed-circuit television with conventionally taught classes in Freshman English both give evidence that television instruction actually is more advantageous to the student of *higher* initial ability. Buckler (2) found in one case that television groups at all three levels of "initial knowledge" improved more than control classes conventionally taught but the differences were significant only in the *high* initial-ability subgroup. In another case conventionally taught groups learned more at all three initial-ability levels but no differences were significant. These groups were tested for gains in standardized tests of English composition. In theme writing the results are somewhat inconsistent: In one case low initial-ability subgroups gained significantly more from conventional instruction than did other low-ability groups taught by television; in the other case high initial-ability groups getting conventional instruction gained significantly more than did a comparable group getting television instruction.

The other study of Freshman English instruction is reported by Seibert (6), who used 10 objective indices of student achievement in a comparison of students taught by closed-circuit television and in the conventional classroom. In one test (spelling) a high-ability television group performed slightly better than the high-ability conventionally taught group, the two middle-ability groups were not markedly different in performance, and in the low-ability groups there was a large discrepancy favoring the conventionally

taught classroom. While none of the other tests was as clear cut as this one, Seibert reports that in six out of the remaining nine tests the same pattern was shown.

Seibert also reports two studies comparing closed-circuit with conventional instruction in mathematics courses under conditions permitting a differentiation of ability levels. In a study of a calculus course (5) no clear patterns emerge; if anything, it appears that conventional instruction is most advantageous at the highest and lowest levels of ability, middle-ability levels showing no difference. However, in a freshman mathematics course subsequently studied (7), a very clear pattern emerges suggesting that television instruction consistently advantaged the high-ability group. Ability was estimated by means of a mathematics placement examination. On six achievement tests the TV group gained higher scores in every instance than the conventionally taught subjects among high-ability students; among the low-ability subjects the TV treatment produced higher scores in four of six tests. But when the magnitude of these differences is examined, it turns out that on all six tests the TV group scores exceed the conventional classroom scores in the high-ability category by a greater difference than that between TV and conventional scores in the low-ability category.[1]

On the other hand, Siegel and Macomber (8) failed to find any consistent differences among ability levels (intelligence) in their experiments comparing TV instruction with small and large classes.

In all the studies cited above, the comparison was between some machine-interposed instructional medium involving both auditory and visual dimensions—whether TV or sound film—with some form of face-to-face instruction. It might be reasonable to ask whether this effect is related to the intervention of the machine and the destruction of the face-to-face contact or whether it may be attributed to the audio-visual nature of the medium. Characteristically the teacher confronted with doing a television version of a course immediately starts looking for "good visuals." There is a lot of face validity to the original hypothesis (that audio-visual media advantage the low-ability student); common sense suggests

[1] Thus $(TV_{Hi} - Con_{Hi}) > (TV_{Lo} - Con_{Lo})$ in all six tests. These observations are based on Table 3, p. 10 of Seibert's report. This table also shows when the six tests are summed, the high-ability TV group exceeds the high-ability conventional group by 465.94 to 434.60, whereas the low-ability conventional group exceeds the low-ability TV group by 411.89 to 400.49. Seibert does not attach importance to these findings "since the design of the study does not include provision for testing the significance of the observed differences." (7, p. 11) However, application of a simple sign test to the published data, e.g., that of Walker and Lev (10), indicates a probability of less than .02 that such consistent differences might be attributable to chance.

that adding a picture may somehow help the duller student more than it does the bright student.

On the latter point it would appear that what is needed is some direct comparison *between* machine-interposed media in which an audio-visual medium is directly compared with an audio-only medium. Just such a study is reported by Williams, Paul, and Ogilvie (11), who compared the effectiveness of four media in conveying the information contained in a verbal message—a college-level lecture on language and culture. On an objective examination immediately following exposure to the message, a television group got significantly higher scores than a radio group. (The TV and radio groups were significantly superior to both a lecture group and a group that simply read the message. Lecture and reading groups were not significantly different.) Eight months later a repeat of the test showed that the rate of forgetting was more or less independent of media. But when the subjects were divided according to general ability based on college grade-point averages, the largest discrepancy between TV and radio scores was at the high-ability level. This difference was not significant, however.

THE STUDY

An opportunity to make a test of the relation between ability levels and learning from television and radio was presented in connection with a study of the effectiveness of radio and television versions of a background-of-the-news program for sixth-graders called "Exploring the News."[2] As reported elsewhere (1), this study found that television subjects learned significantly more information than radio subjects, although the differences were no longer significant six weeks later. It also found evidence that the portions of the broadcast series that made use of the video dimension of television were more salient for the TV subjects than those that did not.

The subjects were all pupils in eight sixth-grade classrooms in four Madison, Wisconsin, public schools (n=228). Subjects were assigned to radio and TV treatments randomly after stratification for verbal and nonverbal intelligence as measured by the California Test of Mental Maturity. The stratification was based on scores earned in an earlier test administration. The analysis, however, is

[2] This study was conducted for the Television Laboratory of the University of Wisconsin with the collaboration of the School of the Air, with the help of partial support from the Educational Television and Radio Center. The cooperation of the staffs of WHA and WHA-TV and of the Madison Public Schools is gratefully acknowledged.

based on a re-examination which occurred after the experiment.[3]

Four 15-minute programs were broadcast on successive Thursday mornings in January 1958, the live television version and the taped radio version being broadcast simultaneously. The two versions presented the same actors in the same roles and were written by the same script-writer. Unlike the study reported by Williams et al, however, who deliberately minimized the possible contribution of the video dimension, our script-writer was under instructions to make each version maximize the effectiveness of that medium. The program content was not identical but its equivalence was empirically established.

The experimental design included several variables other than the ones involved in this report; therefore, data were not available for all four programs from all eight classrooms. However, the data below are based on tests of achievement from one school for each of the four programs, and the assignment of schools to programs was based on chance. (For information on the total design, the earlier article should be consulted.)

An achievement test consisting of 24 multiple-choice information items was administered immediately after each program. Care was taken to assure that the information on which each item was based was presented in substantially the same verbal form in both the radio and television versions. Pretesting was obviated by the highly topical nature of the programs.[4]

Six weeks after the last broadcast in the series a 32-item test was administered to all groups. It consisted of the eight questions from each of the immediate-recall tests which proved by item analysis to be the most discriminating.

RESULTS

Immediate Recall

Subjects were divided into high, medium, and low intelligence scores, high comprising scores above 116, medium 107-116, and low including scores below 107. Table 1 shows the means and cell sizes for scores on the 24-item immediate-recall tests tabled by

[3] Intelligence data used in the stratification were based on tests administered in the third grade. Shortly after this experiment all sixth-graders were re-examined. The more recent data were used in the analysis.

[4] Each program provides background information concerning one or two recent and significant news events. The four programs concerned the presidency and the presidential succession, the opening of the 85th Congress, the Fuchs expedition to the South Pole, and political developments in the Caribbean area.

TABLE 1.—MEAN SCORES ON IMMEDIATE-RECALL TESTS BY TOTAL INTELLIGENCE LEVELS

Total Intelligence	*Radio*		*Television*	
	Means	*n*	*Means*	*n*
High	16.98	43	18.94	32
Medium	15.47	30	16.25	40
Low	13.18	38	15.18	38

intelligence levels and media. Figure I presents the same data graphically. An analysis of variance using the corrections for disproportionality suggested by Snedecor (9) was applied to these data. This analysis is shown in Table 2.

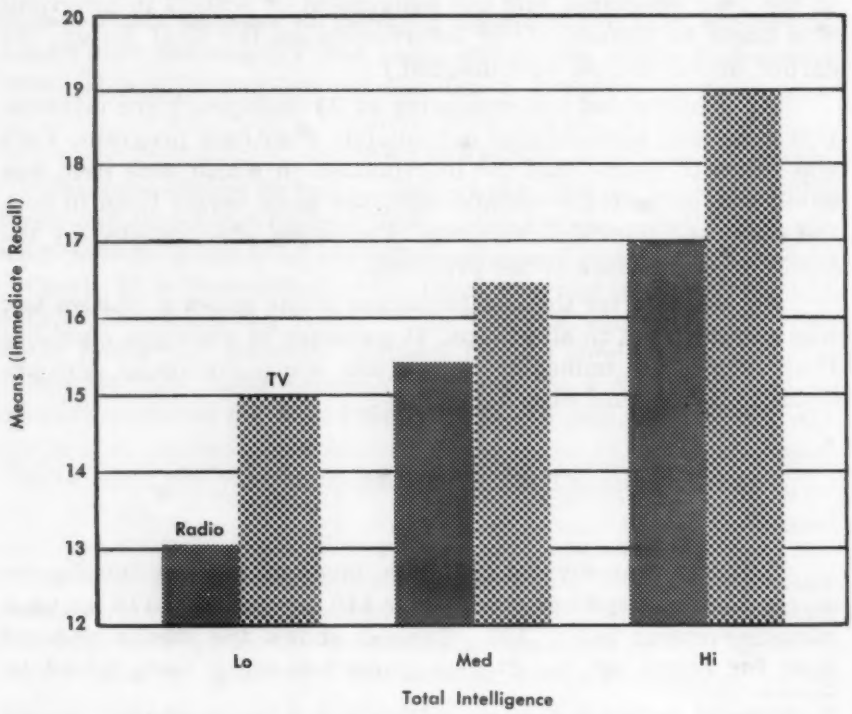


Figure I

Immediate-Recall Means for Radio and Television Groups by Intelligence Levels

TABLE 2.—ANALYSIS OF VARIANCE OF IMMEDIATE-RECALL DATA, TOTAL INTELLIGENCE BY MEDIA (RADIO, TELEVISION)

Source	*df*	*SS*	*MS*	*F*
Media	1	140.19	140.19	9.10*
Intelligence	2	531.19	265.59	17.25**
Interaction	2	16.86	8.43	.547
Within	215	3181.00	15.40	

* *F* significant at .01
** *F* significant at .001

Table 2 shows that there is a highly significant relation between intelligence and the test scores and a highly significant relation between media and the test scores—but that there is no significant interaction between media and intelligence.

However, the question originally raised was whether there is a selective tendency for the media difference to operate differently at various intelligence levels. In order to examine this question, we must test for the significance of differences between means. This analysis is shown in Table 3.

The data in Table 3 suggest that, although we have not obtained a significant interaction between intelligence levels and media, there are indications that the TV-radio difference is operating *at both ends* of the intelligence continuum and not in the middle range. It should be noted that the radio-TV difference is significant at the low and high intelligence levels and not at the medium level. There is further evidence for this in the fact that the radio group means rise significantly between low and medium intelligence levels, but not significantly between medium and high intelligence levels; correspondingly, the TV group means rise significantly between

TABLE 3.—*T*-TESTS FOR SIGNIFICANCE OF DIFFERENCES BETWEEN MEANS, IMMEDIATE RECALL, TOTAL INTELLIGENCE BY MEDIA

Means	*t-value*	*df*	*Significance*
Radio Hi vs. TV Hi	2.18	73	.05
Radio Med vs. TV Med	.84	68	n.s.
Radio Lo vs. TV Lo	2.27	74	.05
Radio Hi vs. Radio Med	1.65	71	n.s.
Radio Hi vs. Radio Lo	4.43	79	.001
Radio Med vs. Radio Lo	2.44	66	.01
TV Hi vs. TV Med	2.94	70	.01
TV Hi vs. TV Lo	4.07	68	.001
TV Med vs. TV Lo	1.23	76	n.s.

medium and high intelligence levels but not significantly between low and medium intelligence levels.

All of this tends to suggest that the narrowing of the differences between media at the medium intelligence levels and the widening of the differences at high and low levels may be real and not chance differences. However, the failure of the interaction between media and intelligence to attain significance must be taken as contrary evidence.

Delayed Recall

Table 4 shows the means and cell sizes for scores on the 32-item delayed-recall test tabled by intelligence levels and media. Figure II presents the same data graphically. The analysis of variance is shown in Table 5.

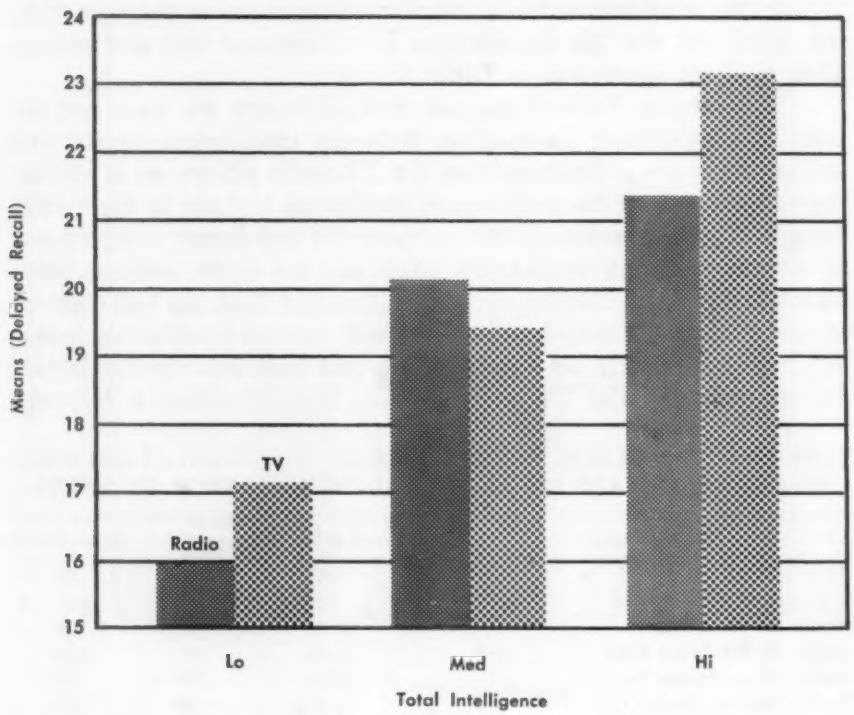


Figure II

Delayed-Recall Means for Radio and Television Groups by Intelligence Levels

TABLE 4.—MEAN SCORES ON DELAYED-RECALL TEST BY TOTAL INTELLIGENCE

Total Intelligence	*Radio*		*Television*	
	Means	*n*	*Means*	*n*
High	21.50	40	23.10	29
Medium	20.19	26	19.66	35
Low	16.07	28	17.40	30

It will be noted that the differences still favor the television treatment at the high and low intelligence levels but that on the delayed-recall test the radio group actually got slightly higher scores than the TV group.

TABLE 5.—ANALYSIS OF VARIANCE OF DELAYED-RECALL DATA, TOTAL INTELLIGENCE BY MEDIA (RADIO, TELEVISION)

Source	*df*	*SS*	*MS*	*F*
Media	1	31.6	31.60	1.69
Intelligence	2	926.6	463.30	24.82*
Interaction	2	41.1	20.55	1.10
Within	182	3398.2	18.67	

* *F* significant at .001

T-tests for the significance of differences between means for the delayed-recall data are shown in Table 6.

The delayed-recall data indicate that the differences between media are no longer significant after the passage of six weeks

TABLE 6.—*T*-TESTS FOR SIGNIFICANCE OF DIFFERENCES BETWEEN DELAYED-RECALL MEANS, TOTAL INTELLIGENCE BY MEDIA

Means	*t-value*	*df*	*Significance*
Radio Hi vs. TV Hi	1.52	67	n.s.
Radio Med vs. TV Med	*	59	n.s.
Radio Lo vs. TV Lo	*	56	n.s.
Radio Hi vs. Radio Med	*	64	n.s.
Radio Hi vs. Radio Lo	5.10	66	.001
Radio Med vs. Radio Lo	3.50	52	.001
TV Hi vs. TV Med	3.19	62	.01
TV Hi vs. TV Lo	5.07	57	.001
TV Med vs. TV Lo	2.10	63	.05

* Less than 1.0

or more. However, the intelligence variable is still highly significant: the value of the *F* has actually increased.

(Direct comparisons between means in Tables 1 and 4 are not possible because they are based on different numbers of items. The smaller *ns* involved in Table 4 reflect the usual attrition.)

Even though television's superiority over radio is still apparent at the high and low intelligence levels, the differences are no longer significant.

DISCUSSION

Whereas previous investigators have found that the advantage of an audio-visual medium over other methods of instruction—when, indeed, it has been found at all—has in some cases been at lower ability levels and in other cases at higher ability levels, our findings appear to suggest that when an unselected population is tested, *the advantage to TV appears to lie at both the high and low levels of intelligence, with less difference appearing in the middle intelligence range.*

It is possible that the explanation may lie in the nature of the populations selected for test. Although this reasoning does not hold for all the data treated in this report, it does appear to apply to most of it. The studies presenting the most convincing evidence that television advantages the *high*-ability group—Seibert (7); Williams, Paul, and Ogilvie (11)—were both conducted with populations restricted at the low-ability end of the range. Both were college populations self-selected by interests and abilities. None of Seibert's "low-ability" group actually received scores on a mathematics orientation test lower than the 84th percentile on the basis of Purdue freshman norms. The most convincing evidence that an audio-visual medium favors the low-ability group was obtained—Kanner, Runyon, and Desiderato (4)—with a military trainee population, which may tend to be restricted at the upper end of the intelligence range.

The present investigation was based on a wholly unselected population. With one minor exception, participants in all cases were the total enrollment in the sixth grades of the four schools concerned, all of which are located in a part of the city which appears to have no particular social class characteristic. The intelligence data were approximately normally distributed, although the means may be slightly higher than a general population average.

It is argued, therefore, that the genotype underlying the relationship between media and learning ability may very well be the curvilinear relationship for which we have found evidence in this study. Further investigation is needed, of course.

VERBAL AND NONVERBAL INTELLIGENCE

The original design made one further provision for examination of the relationship between intelligence and learning from audio-visual versus audio-only media. The California MM embodies verbal and nonverbal subscores. Since an audio-visual medium might be assumed to depend less than an audio-only medium on verbal skills alone, it was reasoned that more powerful predictions of learning levels might be possible from a knowledge of the subject's verbal and nonverbal intelligence levels than on a total intelligence score alone. For example, other things equal, it might be expected that subjects with relatively high verbal intelligence but relatively low nonverbal intelligence would learn more from radio than subjects with relatively high nonverbal intelligence and relatively low verbal intelligence.

All subjects for whom immediate-recall data were available were divided into four intelligence groups: The *Hi-Hi* group consisted of those whose verbal and nonverbal intelligence subscores were higher than average for the group. The *Lo-Lo* group consisted of those whose verbal and nonverbal subscores were lower than average. A *Lo-Hi* group consisted of those whose verbal subscores were lower than average but whose nonverbal subscores were higher than average. A *Hi-Lo* group consisted of those whose verbal subscores were higher than average but whose nonverbal subscores were lower than average.

RESULTS

Means and cell sizes for the 221 pupils so classified are shown in Table 7, together with *t*-tests for the significance of differences between means. The analysis of variance for these data is shown in Table 8.

Once again we find significant *F* values for intelligence and media differences. However, the significant *F* value for intelligence does not provide support for the basic hypothesis. It merely tells us that another way of classifying intelligence again gives us a significant relationship between that variable and learning.

Once again the power of the media difference is demonstrated: all the TV means exceed all the radio means. But actually there is little evidence here for our hypothesis. The only significant difference is irrelevant. Inspection of the table does tend to suggest, however, that verbal intelligence has made a larger contribution in both media. (Both hi-lo means exceed the lo-hi means.) It is true that there is a larger difference favoring television in Subgroup II than in Subgroup III; thus when verbal intelligence is making the

TABLE 7.—MEAN SCORES ON IMMEDIATE-RECALL TESTS BY VERBAL-NONVERBAL INTELLIGENCE CLASSES

Intelligence Subclass (Verbal-Nonverbal)	*Radio*		*Television*		*Mean Differences*
	Means	*n*	*Means*	*n*	
I Hi-Hi	16.89	37	18.78	36	1.886*
II Lo-Hi	14.20	20	15.94	16	1.738
III Hi-Lo	16.28	18	17.60	20	1.322
IV Lo-Lo	13.69	36	14.47	38	.779

* Significant by *t*-test at .05

larger contribution to total intelligence, the superiority of television to radio is greater than when nonverbal intelligence is making the larger contribution. The difference is minute, however, and neither mean difference is significant at .05.

To increase the sensitivity of the verbal-nonverbal component of the intelligence scores a further analysis was made. This was done by first combining the groups in which verbal intelligence is high (I and III) and the groups in which verbal intelligence is low

TABLE 8.—ANALYSIS OF VARIANCE OF IMMEDIATE-RECALL DATA, VERBAL-NONVERBAL INTELLIGENCE BY MEDIA (RADIO, TELEVISION)

Source	*df*	*SS*	*MS*	*F*
Media	1	107	107.00	6.76*
Intelligence	3	585	195.00	3.88**
Interaction	3	12.5	4.17	.281
Within	213	3160	14.84	

* *F* significant at .001
** *F* significant at .01

(II and IV), then recombining to put high and low nonverbal groups together (I and II, III and IV). The means thus obtained are shown in Tables 9 and 10. Results of *t*-tests for significance of differences between means are shown in the same tables.

We note first that when verbal intelligence is varied, all the high-low differences are highly significant. When nonverbal intelligence is varied, *only the TV high-low difference is significant.*[5] This is entirely in line with the hypothesis.

TABLE 9.—MEAN SCORES ON IMMEDIATE-RECALL TEST BY SUBGROUPS REARRANGED TO MAXIMIZE DIFFERENCES IN VERBAL INTELLIGENCE

Verbal Intelligence	*Radio*	*Television*	*Total*	*Mean Diffs (Radio-TV)*
Hi (I + III)	16.69	18.36	17.53	1.666*
Lo (II + IV)	13.88	14.91	14.38	1.032
Total	15.27	16.66	15.96	1.393
Mean Diffs (Hi-Lo)	2.816**	3.450**	3.150**	

* Significant at .05 by *t*-test
** Significant at .001 by *t*-test

TABLE 10.—MEAN SCORES ON IMMEDIATE-RECALL TEST BY SUBGROUPS REARRANGED TO MAXIMIZE DIFFERENCES IN NONVERBAL INTELLIGENCE

Nonverbal Intelligence	*Radio*	*Television*	*Total*	*Mean Diffs (Radio-TV)*
Hi (I + II)	15.95	17.90	16.88	1.956*
Lo (III + IV)	14.56	15.55	15.07	.996
Mean Diffs (Hi-Lo)	1.392	2.352**	1.809**	

* Significant at .05 by *t*-test
** Significant at .01 by *t*-test

There is further support here for the observation that verbal intelligence is making the greater contribution across the board in the fact that the total intelligence differences in Table 9 are considerably greater than the total intelligence differences in Table 10; both, however, are highly significant.

[5] The reader is warned, however, of the pitfalls in interpreting two differences as significantly different when one proves to be significant and the other nonsignificant. The *t*-tests reported here have only ascertained the chances that *each* of these differences is real. This does not entirely satisfy the conditions for a test of the significance of a difference between differences.

TABLE 11.—MEAN SCORES ON DELAYED-RECALL TESTS BY VERBAL-NONVERBAL INTELLIGENCE CLASSES

Intelligence Subclass (Verbal-Nonverbal)	*Radio*		*Television*		*Mean Differences*
	Means	*n*	*Means*	*n*	
I Hi-Hi	21.09	34	22.88	32	1.787
II Lo-Hi	20.06	18	18.40	15	1.656
III Hi-Lo	22.14	14	21.00	18	1.143
IV Lo-Lo	15.96	28	17.03	29	1.070

Delayed-recall data were analyzed in a similar manner. Means and cell sizes are shown in Table 11, the analysis of variance in Table 12. The "crossover" effect at the middle level of ability observed in the case of the total intelligence data again occurs here in both the radio and TV groups. However, none of the radio-TV differences are significant.

TABLE 12.—ANALYSIS OF VARIANCE OF DELAYED-RECALL DATA, VERBAL-NONVERBAL INTELLIGENCE BY MEDIA (RADIO, TELEVISION)

Source	*df*	*SS*	*MS*	*F*
Media	1	10.5	10.50	.58
Intelligence	3	1022.5	340.83	18.76*
Interaction	3	91.2	30.40	1.67
Within	180	3271	18.17	

* *F* significant at .001

When the groups were recombined as above to maximize the verbal-nonverbal effect, the radio-TV difference observed in the case of the immediate-recall data was no longer present. These data are shown in Tables 13 and 14. None of the radio-TV mean

TABLE 13.—MEAN SCORES ON DELAYED-RECALL TEST BY SUBGROUPS REARRANGED TO MAXIMIZE DIFFERENCES IN VERBAL INTELLIGENCE

Verbal Intelligence	*Radio*	*Television*	*Total*	*Mean Diffs (Radio-TV)*
Hi (I + III)	21.40	22.20	21.80	.804
Lo (II + IV)	17.57	17.50	17.53	.065
Total	19.52	20.00	19.76	.478
Mean Diffs (Hi-Lo)	3.830*	4.700*	4.273*	

* Significant at .001 by *t*-test

TABLE 14.—MEAN SCORES ON DELAYED-RECALL TEST BY SUBGROUPS REARRANGED TO MAXIMIZE DIFFERENCES IN NONVERBAL INTELLIGENCE

Nonverbal Intelligence	*Radio*	*Television*	*Total*	*Mean Diffs (Radio-TV)*
Hi (I + II)	20.73	21.45	21.07	.716
Lo (III + IV)	18.02	18.55	18.30	.529
Mean Diffs (Hi-Lo)	2.707*	2.894*	2.767	

* Significant at .01 by *t*-test

differences are significant. Verbal intelligence appears to be making the greater contribution, but actually all the high-low differences are highly significant. Where delayed recall is concerned, then, both verbal and nonverbal intelligence predict differential recall of factual information from both radio and television and there is no evidence that either contributes more than the other to the media.

SUMMARY AND CONCLUSIONS

The relationship between intelligence as measured by the California Test of Mental Maturity and amount of factual information learned from radio and television versions of the same news-background program was investigated using 228 sixth-grade pupils in four public schools. Four weekly 15-minute programs were presented simultaneously over radio and television to groups randomly assigned after stratification for intelligence.

Both media and total intelligence were highly related to scores on an objective test of factual information administered immediately after the programs. The superiority of television over radio was significant with low and high intelligence groups but not with a group of medium intelligence. However, the interaction between intelligence and media was not significant. On a test taken six weeks after the last program, the middle intelligence radio group excelled the corresponding TV group, but TV groups were superior in both the high and low intelligence groups; none of these differences was significant, however.

When the subjects were classified according to part-scores on verbal and nonverbal intelligence there was some indication that verbal intelligence contributed relatively more to learning from

radio than from television, although the indications are—not surprisingly—that verbal intelligence makes a greater contribution than nonverbal intelligence in this verbal learning task.

REFERENCES

1. BARROW, LIONEL C., JR., and WESTLEY, BRUCE H. "Comparative Teaching Effectiveness of Radio and Television." *Audio-Visual Communication Review* 7:14-23; Winter 1959.
2. BUCKLER, WILLIAM E. "A College English Teacher Looks at Television: Composition." *Journal of Educational Sociology* 31:346-52; May 1958.
3. HOBAN, CHARLES F., JR., and VAN ORMER, EDWARD B. *Instructional Film Research, 1918-1950.* SDC Human Engineering Project 20-E-4. Port Washington, N. Y.: Special Devices Center, 1951.
4. KANNER, J. H.; RUNYON, R. P.; and DESIDERATO, O. *Television in Army Training: Evaluation of Television in Army Basic Training.* Technical Report No. 14. Washington, D. C.: Human Resources Research Office, 1954.
5. SEIBERT, W. F. *A Brief Report and Evaluation of Closed-Circuit Instruction in the First Semester Calculus Course.* Lafayette, Indiana: Audio-Visual Center, Purdue University. 1957. (Mimeo)
6. SEIBERT, W. F. *An Evaluation of Televised Instruction in College English Composition.* TVPR Report No. 5. Lafayette, Indiana: Audio-Visual Center, Purdue University, 1958. (Mimeo)
7. SEIBERT, W. F. *An Evaluation of Televised Instruction in College Freshman Mathematics.* TVPR Report No. 6. Lafayette, Indiana: Audio-Visual Center, Purdue University, 1958. (Mimeo)
8. SIEGEL, LAURENCE, and MACOMBER, F. G. "Comparative Effectiveness of Televised and Large Classes and of Small Sections." *Journal of Educational Psychology* 48:371-82; October 1957.
9. SNEDECOR, GEORGE W. *Statistical Methods.* Ames: Iowa State College Press, 1946.
10. WALKER, HELEN, and LEV, J. *Statistical Inference.* New York: Henry Holt & Co., 1953.
11. WILLIAMS, D. C.; PAUL, J.; and OGILVIE, J. C. "Mass Media, Learning and Retention." *Canadian Journal of Psychology* 11:157-63; 1957.

THE GROUP INTERVIEW AS AN AUDIENCE REACTION MEASURE*

• MALCOLM S. MacLEAN, JR.

• HANS H. TOCH

A technique for obtaining audience reaction data by means of group interviews is described, together with examples of its application to two ETV kinescopes. Malcolm S. MacLean, Jr., is associate professor in the College of Communication Arts, and Hans H. Toch is assistant professor in the Department of Psychology, Michigan State University.

THOSE WHO produce and direct educational television programs necessarily are concerned about their impact. How can one get more than the usual feeble information on audience response without going to a great deal of trouble?

The group interview may help to resolve this problem. With a little care, it can provide a fruitful means for learning how people respond to a program. Furthermore, since the producer and director can easily act as participant-observers in the process, they may more quickly derive benefits than they could through rigorous and extensive research. Most of them will likely find group interview results much easier to integrate into a humanistic orientation than the usual research report. Hearing a middle-aged grocer's wife say she simply could not get the point when the man in the picture was scraping the rock with his knife may suggest more to the TV or film man than a neat table of figures. Those who have resisted behavioral research may find group interviews helpful in bridging the gap to a real understanding of what such researchers might contribute and have been contributing to their field. Work with this relatively informal method may eventually lead to a continuing

* The authors are indebted to George Gallup whose work suggested some of the approaches outlined here, and to Edgar Crane, William Tarrant, James Rae, Donald Kiel, and Donald Montgomery of the College of Communication Arts, Michigan State University, for work on the project.

program of research combining both formal and informal approaches.

Developing a research project for the Educational Television and Radio Center at Michigan State, we conducted group interviews with eight different social groups following their exposure to two ETRC programs. Our aims were to:

1. Test out and develop techniques for the group interview method.
2. Attain richer interpretation of more quantitative techniques used in the study.
3. Help in the development of categories for content analysis by discovering which elements in the form and content of the programs were noted by the audience.
4. Reveal some of the variety and quality in response which might suggest hypotheses that could be more rigorously tested later.

The two programs we studied were "Age of Jackson," a lecture by a priest-historian-professor of St. Louis University, and "Making a Formula," a lecture-discussion-demonstration by a woman doctor and a public health nurse.

The "Jackson" program was a simple lecture with several slides from woodcut portraits, and a map of the U. S., used for illustrations. Father Bannon talked about the North-South, the Missouri Compromise controversy, and the views of important political leaders of that day. He discussed the increase in tariffs and gave a simple example of what this meant to the buyer of shoes. He briefed out the Peggy O'Neil scandal. Finally, he showed and recommended three source books on the period.

In the "Formula" program, set in a "kitchen" with an obviously false window and exterior, the doctor compared breast and bottle feeding, then introduced the nurse. Details of cleaning bottles and nipples, mixing canned milk and syrup, and sterilizing were all demonstrated. Then the doctor took over again to tell why cleanliness is so important.

PROCEDURE

What is a group interview? How is it conducted? Here is a description of what we did.

The interviews were semistructured. Although anything relevant mentioned by any member of the group was followed up, a

standard series of points was covered in every session. The interview always opened with the question, *"Now, what did you think of this last program?"* After general impressions had been explored, the viewers were asked what they had *liked best* about the program and what they had *liked least.* These questions were so phrased that any aspects of the program could be mentioned. Responses, in fact, covered the gamut from subject matter, organization, and performer to parts of the program and details of the presentation. In doing this we departed from the premise that in addition to those aspects of the program which add to or detract from its interest for most people, there are specific features whose effects vary from one subgroup to another. Throughout, panel members were encouraged to respond freely.

The next questions asked in our interview dealt with specific portions of the program. The audience was asked to recall parts of the program they particularly liked and disliked. In addition, in sessions in which a program analyzer had been used, the "high point" and "low point" of the profile were located on an outline of the program prepared for this purpose, and were described to the audience. Members were not told that these represented their own group's highs and lows but were asked to recollect their reactions to them. (Thus we could compare the parts they told us they liked with parts the analyzer showed they liked.)

The next question on the schedule asked about learning: *"Do you think you got anything out of this program?"* This was followed by a question about the performer, such as: *"What did you think of the nurse?"* Any performer characteristic mentioned in response to this question was carefully explored, since it seemed that one of the most important values of interviews might lie in providing insight into the cues viewers use to decide whether they like a performer or not.

After covering parts of the program and aspects of the performer, the interviewer asked about the subject matter, inquiring whether the audience felt that it was a worthwhile subject for educational television. Viewers were next told to imagine that they were "put into a position where it was your job to improve this program." They were asked: *"What do you think would be the first thing you'd do?"* The next question, which ended the interview, asked for judgments on the audience reaction device used in the session. The group interviews were all recorded and transcribed. The following excerpts may provide an idea of the responses.

WHAT THE AUDIENCES LIKED AND DISLIKED

For the "Making a Formula" program, most of the favorable comments dealt with format. Audiences in general liked the idea of demonstration, but did not like some aspects of the one in this show. Other favorable references were made to the subject matter, the props, the setting, and to specific behaviors of the performers ("I liked the way the nurse deliberately or indeliberately [sic] made the mistake in making the formula.") There was considerably more unanimity about what was wrong with the program. These are some typical comments:

> Too much detail . . . something like washing dishes to me.
>
> Instead of filling seven bottles, they could have filled one bottle and said, "Here's the way you fill it, see." They could have condensed the program again into about a 5-minute program instead of taking the complete 30 minutes.
>
> It was boring, because there was so much lifting and placing.
>
> Sumpin that kinda irritated me was sometimes they go into this minute detail about sumpin we already know about, like we were little kids or sumpin.
>
> She mentioned about having your measuring cup at eye level, I believe, about three times.
>
> I think she used too simple things, like setting the kettle off of the stove to cool it.

Here, then, was an aspect of this particular program which was of fairly universal concern. If one were to make a single recommendation with regard to this program, it might be that its level should be raised, possibly through condensation and the elimination of any details that might appear obvious. Other things mentioned by the audience as objectionable features were noise ("the kettles and bottles banging together"), performer characteristics, and camera technique. The audiences made specific suggestions concerning production details. One person observed: "It seems as if (the nurse in the program) wasted time by having to move that dishpan over to the counter. It seems as if that could have gone like most of your programs, like your cooking programs, they have someone who takes it from her, she doesn't have to hop all over the kitchen to get rid of it." Seven respondents felt a baby should have been used in the program. Several viewers noticed poor camera work, suggesting that the operator did not follow the movements of the performer sufficiently fast. Miscellaneous suggestions

included the need for more rehearsal, changes in the backdrop, and the addition of more visual aids.

For the "Age of Jackson" program, there was much interest expressed in the subject matter. The presentation of the subject matter, however, evinced both favorable and unfavorable comments. For instance, whereas several respondents observed that the program was very informative, there were 16 comments relating to the number of facts in the program, all to the effect that there was *too much* information. Eleven objections related to continuity. Five charged that facts were being presented too quickly. The following are typical of these types of comments:

> Maybe I'm just not able to take those facts all at one time. It's all right to have the facts, but I couldn't remember them as he gave them.
>
> All those names. He had all those names there. A lot of names. He gave a lot of . . .
>
> Although everything was related, he seemed to go from one subject to another; I mean, he jumped. You sort of wondered how he gets to tariffs after he was giving you a whole list of people and everything.

With regard to visual aids, the picture was similar. Several respondents were favorably impressed by the maps and the slide or two flashed during the program, but a surprisingly large number (20) suggested the use of more visual aids. The consensus was that straight lecturing made the program relatively monotonous, and too much like a college lecture. In several instances "You Are There" was mentioned as an illustration of what could be done to bring history to life.

In both programs, poor camera work also came in for its share of comment. It was especially noticed that there was a lag in following the actions of the performer.

> I didn't like the camera work at all. The man who handled it had it directed at their noses. After a while you couldn't know if they were human beings.
>
> She would show the brush and the camera was focused on the bottle. They weren't coordinated.
>
> Too late in coming in and showing the equipment she was working on.

One general observation possible from a look at some of the preceding comments is that the danger of too much or too little information is a very real one in educational television. Surely such audience reaction sessions as described here might help in arriving at a happy medium between oversaturation and insults to the audience's intelligence. It is also obvious from the above that many subjects mention production details when asked to explain

why they like or dislike a given program. Such flaws do not go unnoticed.

HIGH AND LOW POINTS IN THE PROGRAM

The afore-mentioned likes and dislikes are probably of most interest in relation to a content analysis of programs. It might be possible to define some of the items mentioned objectively enough to be able to determine them directly from the program. Another aspect of the interviews relates to audience reaction profiles. These are the references to specific parts of the programs which constitute "high points" and "low points" to the audience. Again, of course, it might be possible to specify these objectively if their characteristics were known. First, however, we must determine whether the parts which are mentioned in retrospect are indeed the parts which are favorably or unfavorably reacted to while viewing. This is not an easy task in the case of the two programs used in our study, since it developed that there were no real salient features in them. Not only were the profiles singularly flat, but the audience appeared to have difficulty putting their fingers on anything noteworthy. The "Age of Jackson" did appear to contain moments which stood out favorably in the eyes of the audience. Those repeatedly mentioned spontaneously in the interview were:

The introduction with its music (Yankee Doodle)

The discussion of the Missouri Compromise: Thought he did a good job . . . in setting up how the slave states originated. . . . He was touching on a very important subject—segregation—very timely. . . . When he drew lines. . . . When he was talking about the slave states and the free states.

The discussion of protective tariffs, using an illustration of a shoe manufacturer: Where he was talking about the shoes and I began to realize what his point was when he explained cost. . . . Using the board to explain it, about the shoes, I thought that was interesting.

A reference to the Peggy O'Neil scandal: That was interesting to me because I wondered whether it had any connection with the song, "Peggy O'Neil." . . . He did very well with just a little ordinary humor in it, he got our interest, like for instance when he started to talk about Peggy O'Neil.

A listing of possible readings toward the end of the program. Several members of the audience mentioned that they liked the part where "he suggested different books." "I liked the idea of bringing out the actual book and showing it," said one respondent. Another indicated that she "took down all the names of those books and when I go to the library I will proceed to get some." This reaction suggests that group interviews may be helpful in gauging experimental techniques in television production.

Turning to our audience reaction profiles we discover that in general the points discussed spontaneously in the group interviews to explain likes and dislikes correspond closely to those emerging in the program at points the profiles show were liked or disliked. Or, contrariwise, high and low points do acquire psychological meaning in terms of the interview comments. It is apparent that the interviews constitute an important link between reaction measures and content analysis categories, since they suggest *why* our audience liked or disliked what they liked or disliked.

JUDGING THE PERFORMER

Undoubtedly one of the most important determinants of audience reaction to television programs is the audience's reaction to the performer. Common sense and research evidence suggest that if a performer is disliked, the program will be disliked. This makes it very important to know the criteria used by audiences in judging performers, and the determinants of their likes and dislikes. With this information it would not only be possible to select the most promising candidates among potential instructors, but also to advise performers as to those actions and mannerisms which are likely to antagonize or attract viewers.

An examination of relevant comments by our audience shows a tremendous variety of responses. They range from personality evaluations and estimates of personality characteristics (which it might be possible to specify further through intensive interviewing) to rather minute behaviors such as looking at the camera and committing certain types of speech errors. We might illustrate the range of such comments best by quoting some of the phrases actually used. The following are some of the comments about the nurse in the Baby Formula program:

> Interested in her job; didn't have enough interest behind her talks; could have had more facial expression, more hand expression; her voice on same tone; explained . . . very clearly; typical public service nurse; scared; bad grammar; unqualified, nervous, tense, and cold; knew her subject; would have come into your house and helped you out; we all knew she was a good nurse; human.

The public health doctor in the same program was referred to as follows:

> Didn't really think she had something to put over; really didn't care; says "and a"; hasn't got enough talent; personable; very professional; seemed to have everything organized; interested in her job; interesting to

talk to; sort of simple; pleasant person; rather snobbish; capable; mannerisms weren't attractive; talked all on the same level; never was a mother; found her charming; interrupted the nurse; unpleasant, tonal quality; hovered over the other person; without expression; scared; at ease; had more experience; knew her subject; relaxed; tidy; nice smile; unimpressive; pleasant personality; not professional; steady monotone; talked too much.

The most recurrent favorable comment about the performer in "The Age of Jackson" was that he "knew his subject" or "knew what he was talking about." Eleven comments took this form. Other favorable comments were:

Relaxed; clear voice; professional; didn't talk over anyone's head; enjoyed what he was doing; had life to him; warm personality; I thought he was very—he was—personality; natural; he really meant what he said; I like the way he dressed; much like Bishop Sheen; he held my interest but I don't know why; he was himself; he was not an actor; it wasn't mechanical; typical good educator; actually talking in conversation; I felt as though he was in person; clear; audible and easily understood; very pleasant voice; stumbling only natural; he was looking at us; not wandering all over the place; human; right in the room; he was the instructor and we were in class; realistic; I liked his easy conversation; you could like very much, how he said "Hi!" as if he knew you; he could talk to the class all day and we would sit right there and listen to him; he went along with the subject matter with his facial expressions; sincere; he made it so much a living thing; he used gestures; I wanted to listen; he got your interest; magnetic; marvelous teacher.

The most frequent unfavorable comment about the same performer was the charge that he stuttered and stammered (nine mentions). There were also five instances in which viewers expressed a reservation about a priest teaching history. Other unfavorable comments were:

Lacked personality; I'd rather have a woman teacher for history; too serious; wasn't congenial acting; no facial expression; amateur; cold; not at ease; too many ah's and a's; moved around too much; lacked preparation; floundered around too much; kept his hands going too much; not relaxed; poor delivery; jumpy.

A rich and interesting area for further work appears to lie here. First, we can try to specify in a content code the characteristics of performance, speech, et cetera, which audiences take into account in reacting to a performer. Second, we can explore an audience's bases for such complex judgments as whether a performer knows his material, whether he is interested in people, whether he is friendly and congenial, or whether he has that mysterious set of qualities people refer to as "personality."

LIMITATIONS

While much of the material from group interviews may provide valuable illumination and suggest fruitful hypotheses, it can seldom be used for testing hypotheses or "proving out" any general principles.

Dominating, talkative members of an audience panel, unless special care is taken, may tend to discourage expressions of opinions divergent from their own. This may lead to a biased picture of audience reactions—we may get only a minority view—especially where the group interview method is not supplemented by others which allow a person to say privately how he feels.

Sampling problems arise, too. But these occur in any study where one must ask a small cross section of his potential viewers to come out of their homes, sometimes in unpleasant weather, ride in car or bus to a place he can use for the viewing session, and sit and view something which they ordinarily might not have chosen to view.

SUGGESTIONS ON TECHNIQUES

Line of Questioning

Where the producers and directors of films or programs can spell out the objectives of their productions, development of meaningful interview questions becomes much easier. General areas for consideration might include these:

A. Attention
   - I. What kinds of people will your program appeal to?
   - II. In what ways, if any, do audience members actually find it rewarding?
   - III. If this production is one in a series, will attention to this lead to the desire to view others in the series?

B. Understanding
   - I. Given a particular production, how much does it contribute to an increased comprehension of the subject matter with which it deals?
   - II. For its intended audience, is its complexity level too high or too low?
   - III. Within the production, which things seem to clarify, which to muddy or befuddle? What is found relevant, what irrelevant?

C. Attitude

I. Some productions aim to reinforce or modify the attitudes of their intended audiences. Do the feelings expressed by viewing panel members toward the subject of the program or film reflect those desired by its creators?

II. Productions may aim to develop the "self image" of the person who views. Does he feel that the particular production has done something for him, increased his self-assurance, given him the impression that he can be more worthwhile, expanded his view of the universe, raised his joy in living, et cetera? Has it given him new ways of looking at the people and things in the world around him? Has it even changed his "mood"?

D. Action

I. A production may attempt to get its viewer to do something, to read a book or paint a picture or form an action group, for example. Is he aware of what it wants him to do?

II. If he is aware, does he agree with the idea? Does he feel it is realistic to expect him to carry out the proposed action? How high does he estimate the probability that he will carry it out?

Sampling

Whose reactions should be studied? When those who conduct the study know at whom the production is aimed, they can draw an appropriate sample. For example, a sample of student and practicing nurses might be drawn to view a film dealing with nursing techniques. The researchers might select actual and potential union business agents, arbitration and mediation specialists, and industrial and business executives to view a television program on labor-management relations. When a program is intended for broadcast to the general adult audience, they might wish to get a representative sample of that audience.

One should try to imagine all the various kinds of people for whom the program or film might be useful. The labor-management program possibly is salient not only for the people mentioned, but also for their wives, for union members, supervisors and foremen, for college and high-school students in social science, and so on—

depending on what aspects of labor-management relations are emphasized. It is probably less costly to get more groups and find a few who do not see the production as relevant than to neglect some who might find it quite meaningful.

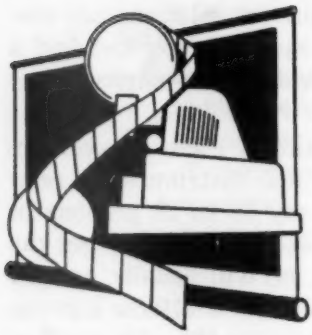
Since production people primarily want to learn the variety and quality of reactions, not their statistical distributions, relatively small viewer panels should suffice—say 10 to 20 persons in each group. While the study director can often more easily get the cooperation of an intact, organized group, he will usually not get the variety of response from such a group that he will from a group made up of one or two individuals selected from each of a number of intact groups. Generally speaking, it is better to take one or two members from a number of different union locals than all from the same local.

Will the sample be seriously biased since it only includes people from one community? Very likely some varieties of reaction will not be represented. But the researcher can use his knowledge of the community, how it compares with others where the production might be shown, to interpret the group interviews. In any case, these interviews provide suggestive materials, not conclusive evidence.

To obtain a group from the general population one might start with a large sample from the city or telephone directory, expecting that many persons cannot or will not come. Or, if program objectives suggest specific groups, membership lists of those groups may be used, particular members being selected by some unbiased system.

Aside from the audience information obtained, considerable publicity benefits may develop from this informal research activity. People dispersed through the whole community, in many different occupations, get to know the production unit and what it is doing.

Group interviews, in short, can serve several purposes, or combinations of purposes. They provide a relatively easy, inexpensive way of getting audience reaction data. They yield rich, quotable human interest material that appeals to people who are left cold or worse by statistical tables. They highlight the qualitative aspects of results obtained through more rigorous assessment techniques. Lastly, but not least, group interview sessions serve a two-way public relations function. They bring the producer and his audience closer together than other research techniques can do.



FILM

• CAROLYN GUSS

FILM EVALUATION PROCEDURE USED AT INDIANA UNIVERSITY AS APPLIED TO A REPRESENTATIVE SELECTION OF RECENT EDUCATIONAL MOTION PICTURES

Since it is neither desirable nor practicable for a single film library to circulate all available films or for a user to use all available films on a given subject, evaluation and selection must take place. Within the last 10 years, the number of educational film titles has more than tripled—the 1953 and the 1954-1958 *Educational Film Guides* list and describe over 17,000 educational motion pictures. The increased activity in film production has made available several different motion pictures on the same subject for the same grade level. Whereas not so many years ago a film librarian or potential user was content to be able to locate one film title on a given subject and grade level, today he is frequently bewildered by finding five, six, seven, or more films covering essentially the same content.

Evaluative criteria or standards of excellence must be used consciously or unconsciously by the evaluator in arriving at an estimate of the worth of the film in terms of the purposes which he has in mind. These criteria may have a great deal of specificity or they may be very general: they may be written or not; they may be weighted in terms of their importance, or their relative importance may not be indicated. As Charles F. Hoban, Jr., has said, "The individual in charge of a film library selects films from the total quantity available, and the teacher selects from the library particular films for use. Whenever there is selection of films, there are standards of selection. Sometimes these standards are explicit, at

Carolyn Guss is associate in selection in the Audio-Visual Center and associate professor of education in the School of Education, Indiana University.

other times they are acted upon without recognition: but always there are standards."

Among the standards used by Indiana University faculty members and graduate assistants evaluating new educational films are the following:

A. General factors
   1. Does the film serve a socially desirable and educationally important purpose?
   2. Do the component parts of the motion picture complement and supplement each other in such a fashion as to result in a film which gives a total impression of unity and a satisfactory synthesis?
   3. Are the range of material and the range of appeal to maturity level compatible with using the film at a given grade level for a specified purpose?

B. Content factors
   1. Is the main idea in the film developed in a constellation of ideas which are mutually interbehaving and interacting?
   2. Are the solution to the problem, the development of certain mental and personality behaviors, and/or the acquisition of skills adequately developed?
   3. Are categorical representations sufficiently supported by differentiating details so that those in the audience will not base their generalizations on stereotypes?
   4. Are any misconceptions likely to be formed because of overcondensation?
   5. Are there any serious omissions in the content?
   6. How effective are the positive or negative examples?
   7. Is the treatment of content effective in terms of the purposes of the film?

C. Psychological factors
   1. Does the film provide a set which will give direction to behavior and provide incentives which individuals will strive to attain?
   2. Does the film provide for audience participation—covert as well as overt?

3. Is it conducive to audience identification and ego-involvement?
4. Does the film stimulate interest?

D. Technical factors

1. Is the photographic quality sufficiently satisfactory to enable the audience to see the photographic symbols without undue concentration on the process of seeing?
2. Is the sound easily intelligible?
3. Are adequate orientation devices used?
4. Is the type of pictorial representation appropriate for presentation of the main ideas in the film?
5. Does the type of sound accompaniment make the greatest possible contribution to developing the explanation and providing emotional tone for the main ideas in the film?

During the past year faculty-student groups at Indiana University using these criteria have evaluated over 500 educational films as they were released. Representative of their evaluations are the following 10 condensations.[1] The first paragraph of each gives a brief over-all statement of content and the second, a brief appraisal of usefulness.

Burden of Truth. United Steelworkers of America, 1500 Commonealth Building, Pittsburgh 22, Pennsylvania, 67 minutes, 16mm, sound, b&w, 1957. Apply to distributor for terms governing use. Produced by Allend'or Productions. Teacher's guide available.

This film focuses on the theme of anti-Negro discriminations and segregation in American life and challenges the viewers to accept the burden of truth. It assembles correlated incidents to represent a pattern that operates in the society and through its char-

[1] These condensations are based upon Indiana University faculty-graduate student previews reported on Educational Film Library Association evaluation forms and subsequent longer reviews prepared for *Educational Screen and Audio-Visual Guide*. The following faculty members and graduate assistants in the Audio-Visual Center contributed to these reviews: James W. Armstrong, Jr., Theodore Beers, Henry Bern, O. E. Bissmeyer, Jr., Beryl Blain, Bruce Buckley, Ledford Carter, J. R. Eichorn, Rachel Feldman, Malcolm Fleming, Seymour Friedberg, Roy Frye, Sheldon A. Goodman, Merlyn Herrick, Maurice Kessman, Robert Kinker, James Q. Knowlton, Robert McAdam, Francis Moakley, Jean Holt Moore, Robert B. Pettijohn, Stuart Roe, Warren Stevens, Kenneth B. Thurston, George Vuke, Robert Weisgerber, William A. Wheeler, Fred E. Williams.

acterizations it presents prevailing sentiments and practices along with the usual arguments which are used to justify their existence. Joe Hamilton, a Negro youth, is presented as the central character of the plot and his experiences are dramatized first as a janitor in a saloon in his Southern hometown, then as a working college student in a Northern university, and later as an industrial worker in a Northern city.

Leaders in the fields of human relations or race relations and teachers of social studies will find this film useful in provoking thought and discussion in their groups. They must be on guard, however, to guide the discussion into fruitful channels because the film does not point to specific issues for discussion but, rather, points to general problems. The purpose of the film is to bring to light discriminatory practices which work against certain ethnic minority groups. It should be useful in connection with social problems in groups which range from the high-school level to college and adult levels. The target audience of the film seems to be general adult groups.

City of Gold. McGraw-Hill Text-Films, 330 West 42nd Street, New York 36, New York, 23 minutes, 16mm, sound, b&w, 1957. $130. Produced by the National Film Board of Canada.

City of Gold is about Dawson City, its present and its past, as seen from the viewpoint of narrator Pierre Berton, "who was raised there." It is a quiet ghost town now where three or four hundred hard-working people live. The children playing baseball in the park think of Dawson City as a beginning, but the weathered old men who sit along the main street talk of the "good old days." In the summer of 1898, 40,000 souls were here in this boomtown of the Klondike Gold Rush.

High-school and college history teachers would find the film useful in portraying conditions during the gold rush era. Since the Klondike Gold Rush is directly related to the settlement of the Canadian West, students of Canadian history and historical geography will find the film even more pertinent. When a high-school teacher previews this film, he will decide whether the fashionably dressed "ladies" from "Paradise Alley" and paintings of Victorian nudes will disqualify the film for his use situation, or whether these aspects of the film are "facts of life" in a gold rush community presented subtly and inoffensively to a sophisticated audience.

Communication in the Modern World. Coronet Instructional Films, Coronet Building, Chicago 1, Illinois, 11 minutes, 16mm, sound, color or b&w, 1959. $110 or $60. Teacher's guide available.

This is the story of how ideas and feelings are communicated from one person to another through a variety of media. The narrator explains that there are two forms of communication, sight and sound. Examples of these two types are presented, such as gestures, smoke signals, traffic lights, policeman's whistle, jungle drums, and the telephone. Motion pictures and television evidence a combination of the two.

The evaluating committee feels that *Communication in the Modern World* is especially useful as an introductory film for a unit on communication in the intermediate and junior high grades. As a broad overview of the great variety of communication channels in use, it serves the necessary role of introducing the student to the urgency of having rapid, effective communication and to the number of ways one communicates.

Gateways to the Mind—The Story of the Human Senses. Produced for Bell System, 60 minutes, 16mm, sound, color, 1958. Apply to your nearest Bell Telephone System office for use. Teacher's guide and student's guide available.

This film, the fifth in the Bell System Science Series, uses animated figures, diagrammatical drawings, live action, and documented film footage to dramatize what happens when we hear, see, smell, taste, and feel. Dr. Frank Baxter, in a Hollywood sound stage setting, acts as narrator and guide.

Gateways to the Mind continues the outstanding precedent set by the other films in the Bell System Science Series by having an interesting attention-holding treatment, accurate content, clear presentation, coupled with excellent color, sound, and acting. Dr. Frank Baxter again excellently performs his role as narrator and guide. The two broad purposes of the film may be stated as follows: (1) to present factual information concerning the senses resulting in a more understanding and deeper appreciation of the senses and motivating further study concerning them, and (2) to create a feeling of respect for the work of the scientist and consequently serve to attract students to the field of science.

Give Them a Chance. Pennsylvania State University, University Park, Pennsylvania, 12 minutes, 16mm, sound, b&w, 1957. $60.

This film surveys the curriculum experiences of mentally retarded pupils and shows how the children learn and react in a setting which is adapted to their needs. It explains the objectives of the program for retarded children and points up the values inherent in the types of activities which are afforded these pupils. The role played by a competent and understanding teacher is shown throughout the film by subtle implication. The thesis of the film—that mentally retarded children can succeed if given a chance—is developed.

Give Them a Chance effectively pleads the case for special education of mentally retarded children. It is sympathetic in its treatment but not unduly sentimental. Teachers-in-training and teachers on the job will find this film useful in focusing their attention on the problems of slow learners. Parents can gain some idea of the kinds of experiences that are provided children who can not successfully adjust to the ordinary classroom procedures. The technical quality of this film is not excellent, but its treatment of the subject matter compensates for this inadequacy.

Learning About Flowers. Encyclopaedia Britannica Films, Wilmette, Illinois, 10 minutes, 16mm, sound, color or b&w, 1958. $120 or $60. Teacher's guide available.

This film makes frequent use of time-lapse photography to show the opening of the flower buds on some of the more common cultivated and wild plants and tells of the importance of flowers to the plants and to man.

Even though in most cases the film indicates the approximate time the process observed by time-lapse photography in the film would take in nature, it is suggested that before showing the film the teacher clarify the relationship between actual time and screen time. The personalized narration and the technical excellence contribute much to the film's effectiveness. For example, the narrator suggests without being too obvious or authoritarian that it is better to leave the flowers in the field than to pick them. This film will probably be most useful on the primary and intermediate grade levels.

Pictures Teach at Penfield. Local Eastman Kodak Audio-Visual dealer or Eastman Kodak Company, 400 Plymouth Avenue, N.,

Rochester, New York, 17 minutes, 16mm, sound, color, 1958. Rental free. For purchase information apply to Eastman Kodak Company, producer of the film.

With James Meagher, instructional materials coordinator at Penfield School, New York, as narrator and guide, the film shows through the case-study technique how the audio-visual program in the Penfield schools helps boys and girls learn. It documents with sound photography educational practices in the utilization, distribution, and local preparation of audio-visual materials.

On the basis of several actual uses of this film with professional groups interested in the educational values of the audio-visual program, the evaluating committee highly recommends it as useful for showing (a) the values of an audio-visual program both to students in classes and to the student assistants, (b) the scope and operation of a representative audio-visual program, and (c) the interrelationships and responsibilities of faculty, students, and community as they develop the best possible kind of learning environment and experiences for boys and girls.

Prehistoric Images: The First Art of Man. Brandon Films, Inc., 200 West 57th Street, New York 19, New York, 16 minutes, 16mm, sound, color, 1955. $250.

"The artist records for those who come after the image of the world he knows." Thus, begins a journey backward in time, revisiting the subterranean world of the cave painters of prehistoric Europe. Through the cooperation of the French and Spanish governments, the film recreates for the beholder the rarely seen images traced on the walls of the caverns of Lascaux, Pêch Merle, Niaux, and Trois Frères in France, and Altamira, El Castillo in Spain more than 50 centuries ago, when these early artists sought shelter there from the bitter cold left behind by the retreating ice sheet.

The experience which this film presents is a blend of the art of the movie-maker and of the painter. The story is told with little need of help from the sound track, which is low-key and somewhat hypnotic (the sound quality of the print reviewed was unfortunately poor). Variations in camera movement and pacing, lighting, and optical effects combine to keep the film a fascinating treat to the eye throughout—despite its motionless subject matter—from the grandeur of the cavern to the violence of the imagined chase.

It should be interesting fare to students of art and man, whatever their ages.

Problems of the Middle East. Atlantis Productions, Inc., 7967 Sunset Boulevard, Hollywood 46, California, 21 minutes, 16mm, sound, color or b&w, 1958. $200 or $120.

This film depicts the history and culture of the Middle East area from antiquity to the present resurgence of nationalism and unrest. Its emphasis is on the anthropological approach as a key to the understanding of the varied and complex issues involved in this troubled area. Four main problems are presented and analyzed: the place of minorities, the area's agricultural need and potential, the rise and impact of Western technology and industry, and, lastly, the role of education.

Few subjects are more complex and more inexhaustible than that of this film. Its main purpose, to acquaint the viewer with the historical and anthropological backgrounds of the Middle East, is well realized. Thus the film makes a valuable contribution to senior high and adult education levels. Particularly well drawn out is the present Arab dream of an Arab world stemming from the Islamic empire of the Middle Ages and the unifying factors of a common culture, language, and religion. In contrast, the sequence on the educational needs of the Middle East is most inadequately developed. Perhaps this lack, as well as the too briefly exposed maps, will be cared for in the reading and discussion which surely must follow the showing of this well-conceived and colorful film.

The Return. American Physical Therapy Association, 1790 Broadway, New York 19, New York, 39 minutes, 16mm, sound, b&w, 1958. $140. Produced by MPO Productions, Inc., New York. 28-minute version also available.

The Return is in semidocumentary dramatic form, depicting the role of the physical therapist in the rehabilitation of a paraplegic victim.

The film is designed to be used in vocational guidance programs and will definitely stimulate a desire on the part of the viewer to seek further information about a career as a physical therapist. It may be used profitably by senior high-school and college guidance counselors, in nurses' training programs, and in general adult groups.

BOOK REVIEWS

MAY, MARK A., and LUMSDAINE, ARTHUR. (In collaboration with **HADSELL, R. S.; GLADSTONE, A. I.; and HOWELL, J. J.**) *Learning from Films.* New Haven: Yale University Press, 1958. 351 p.

Reviewed by

• PAUL R. WENDT

This volume represents another large serious step forward in research in audio-visual education. It is a report of the Yale Motion Picture Research Project begun in 1946 with a grant from the Motion Picture Association of America, expanded with support from Teaching Film Custodians, Inc., and concluded in 1954. The senior author was director of the project for the entire period; Arthur A. Lumsdaine was associate director until 1949. Lumsdaine was the author of most of the experimental studies; Mark May authored many of the remaining studies.

The book is organized into four parts after an introduction by Mark May. Part I describes more than 11 experiments, most of which were performed by Lumsdaine, some with the assistance of Gladstone, May, and Hadsell. This section is mainly concerned with production variables in the films themselves and utilization factors. Part II, principally the work of Mark May, is concerned with the effects of films on subsequent learning activities such as reading, participation in class discussions, and changes in attitudes. Part III describes in detail the testing techniques used. Part IV contains two surveys on audio-visual education in Connecticut and the concluding chapter.

Paul R. Wendt is chairman of the Department of Instructional Materials, Southern Illinois University, Carbondale.

Part I is the core of the book since it contains reports of attempts at basic research about films, carefully controlled experimentation, and modern statistical treatment of the data. Each chapter attacks some intriguing variable in the production or use of educational films. Chapter 2, for example, compares the effectiveness of a color-sound film with a "pencil-test" version of the same film. The pencil-test reel was a crude black-and-white motion picture made from the 200 story-board sketches made originally by the producer during the planning stage of the film. This pencil-test reel was synchronized with the planned narration of the film. Therefore, this experiment compared the effectiveness of the complete color-sound film with the palest, crudest version of the film possible. The two versions were shown to four fifth-grade classes divided randomly. A 91-item test revealed no significant differences in learning—a shocking result when we consider one film version probably cost 10 times as much as the other.

This experiment also demonstrates the characteristic analytical hypothesizing of the authors. They speculated that the apparent lack of difference between the two film versions might be caused by each version being especially effective in certain areas of content, the two cancelling each other out in the over-all means. A further analysis comparing the test scores in 10 content areas of the film still revealed no significant differences.

This experiment was followed up by another comparing the effectiveness of the color-sound version with the black-and-white-sound print of the same film. This was shown to both fifth- and ninth-grade classes. No significant differences were found. Likewise, student ratings on a 10-point scale of their interest in the film, interest in the subject matter, and understanding of the subject showed no superiority for either version.

These results are all the more amazing because the film, *The Seasons,* theoretically could have made effective use of color. The de-emphasizing of color, filmic motion, and the general artistic qualities of a teaching film agrees with the findings of Hoban, van Ormer, and Zuckarman. Whatever superior qualities the esthetics of film production contribute are not revealed by achievement tests.

Another experiment compared a version of a film containing considerable dialogue with a version of the same film containing off-stage commentary only. Again there were no significant differences. The above three experiments might indicate to producers that expensive production techniques are not translated into learning values in their films.

Two versions of another test film differed only in the difficulty of the commentary as measured by the Dale-Chall and the Flesch Reading Formulas. Detailed analysis and sub-analysis of the data allowed the experimenter, Gladstone, to make only very cautious conclusions since the effect of the difficult commentary was slight. Some may question his basic premise that readability can predict listenability. The author quotes several studies indicating a fairly high correlation between readability scores and listening comprehension, but a much deeper study needs to be made of this relationship.

A totally different kind of experiment used a slide-film teaching the phonetic alphabet to test two other variables. The first variable was active vs. passive review. In active review the class was asked to participate orally. The second factor was comedy: the pictures on the screen contained cartoons and the commentator made comic remarks. The total experiment was then set up in a simple two-by-two factorial design. Both active review and simple (noncomic) presentations came out superior. An interesting sub-analysis of the ability of the subjects revealed that active review was equally valuable for all levels of ability, but that the noncomic presentation was more important for those of lower ability, another possible example of how careful research can upset "common sense" predictions.

Splicing questions into a film was the basis of another experiment. Questions were spliced into a film before a sequence which answered the questions to improve motivation, and after the sequences in a film to promote participation. Besides the two versions, some students saw the original film and some saw the film with both types of questions spliced in. Motivating questions had no significant effect, but participation questions had a pronounced result. The version containing both motivating and participating questions also had a pronounced result, but analysis showed this to be almost entirely due to the participating questions alone. It might be suggested here that the effectiveness of participating questions depended upon the immediate knowledge of results which accompanied them. In a sense, therefore, we have an example of operant conditioning in action. The statistical analysis of this experiment is particularly interesting.

A penetrating experiment on utilization tested two factors—whether calling attention to certain parts of a film would increase learning of those parts and whether this learning would be at the expense of other parts of the film. The film used was a complex

one, *David Copperfield.* One group of students was asked to pay particular attention to the characters, another to the plot, and a third to the settings. They were also warned that they would be tested on these areas after the film showing. In each case the results were pronounced, thus verifying, in regard to films, an established classroom principle. On the question of whether directing attention to parts of the film decreases the learning from the other parts, the results were inconclusive. This means that at least there were no significant losses in learning in the other parts of the film.

The pre-test on parts of the film was also used to direct attention to these parts—this proved effective. A review after the film showing of the 10 hardest questions followed by a second showing also produced significant gain, again a verification of good classroom teaching method.

A final utilization test consisted of the film showing, a post-test, the papers handed back to each student with the errors plainly marked, followed by a second film showing and a final test. Here, strangely, no significant improvement was recorded. The authors have theories about this negative result—but no answers. It should provide a glittering opportunity for some further experimenters to attack this intriguing problem.

It might be noted that all the above utilization experiments are not indigenous to films alone. Teachers know that directed attention and the threat of post-tests are usually effective for any learning methods or materials. It is good, however, to have these particularly demonstrated with teaching films.

The grade level of a film came in for testing. The film, *The Seasons,* was rated by its producer as a fourth- or fifth-grade film. It was shown to a group of fifth-graders and also ninth-graders. The report here reveals one of the outstanding features of this volume—the probing statistical analysis. The results of the fifth- and ninth-graders were compared in three different ways: (a) by comparing mean gains in the number of items correct; (b) by comparing the net proportion of subjects improving their scores by one or more points; and (c) by using an effectiveness index that expressed mean gain as a percentage of possible gain. The ninth-graders were superior in all three measures. The effectiveness index is a sophisticated technique that should be used more widely.

A sub-analysis investigated the differences between the two grades on the 10 content areas of the film. No outstanding differences were noted. The sub-analysis of the data from a film experi-

ment shows us that real progress has been made in the last two decades in digging deeper into the cause of film effects. Experimentation has come a long way from the day when it was sufficient to test whether a film as an entity produced learning.

Another film, *Osmosis,* was shown at the end of the school year to one group and at the beginning of the school year to another, thus testing the effects of previous instruction in the subject. The end-of-the-year group learned significantly more from the film.

Chapter 10 reports on an entirely different kind of experiment employing the method of paired associates. The pairs consisted of word-word, word-picture, picture-word, and picture-picture. The data from this experiment were worked and reworked to produce as many analyses as possible. The results are complex and cannot be reported here.

The methodology raises a question. In a word-to-word pairing the presentation of the first member of the pair should evoke the second. In a picture-to-word pairing the picture should evoke the paired word. But the authors do not make clear what the response should be in a picture-to-picture or a word-to-picture pair. It seems that the response must be verbal. In this case there is an additional word involved. The two sequences would be picture-picture-word and word-picture-word.

The method of presentation of the stimuli in a classroom as reported states that the image size on a translucent screen was 7 by 11 inches and that some subjects sat as far as 30 feet from the screen. Since well-accepted projection standards in this case would call for a 5-foot screen, there is a question here about the validity of the viewing situation. At least the authors were obligated to acknowledge the situation and offer an explanation, if not a validation.

A further analysis of the experiments with the films, *The Seasons* and *Osmosis,* attempted to test effectiveness of individual sequences in the film as determined by a few items on those sequences in the post-test. The results were inconclusive.

The final chapter in Part I does not report any experiment in the series but is a theoretical discussion of how demonstration films should be planned, based on some psychological principles and some previous experiments.

In Part II the effect of films on subsequent behavior was reported. The showing of the film, *David Copperfield,* gave disappointing results as measured by withdrawal of books from the

library, or stated preference for books to read in or out of class or to receive as a gift.

Stevenson's *Kidnapped* was selected to test whether showing the whole film or only a few exciting episodes would motivate the class to read the book. The excerpts were more effective in getting students to withdraw the book from the library, but there were no differences in the students' reported readings or test scores. Of course, students' reports of their own activities or interests are not always very reliable.

In New Haven, Connecticut, a study was made of classes using films heavily, as against those who did not, compared on the basis of reading ability. Pronounced differences were obtained. The author suspected this might be due to better teachers, but the teacher rating scales they used were inconclusive. Similar negative results on other variables caused the authors to conclude that the films themselves were mainly responsible for the improvement in reading. This important result should be noted. It is hoped that other cities with records of film use will replicate this experiment and strengthen its significance to audio-visual instruction.

Thirty-six class discussions about six films provided data which were very carefully analyzed. This is an area, however, in which intervening variables make experimentation extremely difficult. Even with the best analysis (as in this experiment) control of complex variables is hard, if not impossible. For example, the distribution of class members taking part in discussion is violently skewed. A few students tend to do most of the talking. Experience in discussion techniques is crucial. The authors found the teachers doing more talking than all of the students combined. The results, then, were inconclusive.

A similarly difficult experiment attempted to show a change in attitude toward "due process of law" by the showing of two films. After the usual rigorous statistical analysis, the authors were forced to conclude that this experiment provided all the useful clues for further work in this area.

Part III on the techniques of evaluating instructional films provides two of the most interesting chapters to research workers. The first one on evaluation by objective tests is a scholarly résumé of the methodology used in the Yale Motion Picture Project and contains a discussion of the effectiveness index mentioned above. The chapter on film evaluation by "road-testing" also makes a contribution in the form of a standardized technique for subjectively evaluating an educational motion picture.

Part IV reports two surveys in Connecticut. One is a general survey of audio-visual education in the state and the other is a survey of factors related to the use of motion pictures as reported by 2700 questionnaires sent out to teachers.

Learning from Films, therefore, constitutes a real advance in the science of testing films. The variables tested, especially in Part I, are basic and varied. Outstanding are the methodology and statistical treatment.

The volume is not without faults. Criticism could be centered on two important areas. One might question over and over in this project why attempts were made to prove anything by showing just one film, sometimes only 10 minutes in length. What other activity in the school is expected to produce any results in 10 minutes? The most effective teacher in teaching methods would be hard put to produce demonstrable changes in a week of teaching. Why should the educational motion pictures be singled out to produce instantaneous miracles? Many experimenters in the audio-visual field have been using 5, 10, and even 50 films to show long-range gains. Elaborate statistical techniques cannot squeeze many drops of factual blood out of a one-film turnip.

This deficiency is aggravated by the authors' tendency to over-elaborate discussions of each experiment. They are punctilious about separating their statements of results from their discussions and every reader should also make a clean-cut distinction between these two. Unfortunately there is a danger that casual readers, skimming the volume and looking for conclusions at the end of each chapter, may be carried away by some of the unwarranted theorizing under the discussions. For example, the strong doubts raised about the value of esthetic qualities in the film, about color, about live dialogue, about the difficulty of commentary as measured by readability formulas, the recommendations against comic touches and in favor of spliced-in questions on the screen—too many of these are based on a showing of just *one film.* A good example of this occurs in Chapter 9 where the author says, "Given a choice of whether to use a film of the usual classroom type, used only once in a course, [the writer] would use it as a summary rather than as an introduction." Notice the expression "a film." This makes all teaching films sound like standardized products expounding facts or processes. It ignores the wide range of different types of teaching films for different purposes. The discussion sections of the book are full of such extreme generalizations, unjustified by the evidence brought up by excellent research. Some of the

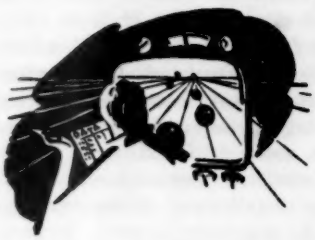
theorizing and discussion sections attempting to grasp the significance of some of the results will provide further researchers with excellent clues, but because of the danger of misleading casual readers, the discussion sections could better have been much more modest.

The second major criticism of the experiments is that so many of them ignore the utilization by the classroom teacher entirely. After Wittich's experiment in 1946, it has been generally recognized that the way a teacher uses a film is far more important than the film itself. In many chapters of this book the classroom teacher is not even mentioned. Undoubtedly the authors controlled this factor, but they had a responsibility to state explicitly how they did so. In the concluding chapter, Mark May intimates that films may be used more extensively as a self-teaching device without any teacher utilization. If this was not merely an afterthought, it should have been explicitly stated in the experimental designs. Of course, some of the experiments were investigating utilization itself.

There might also be some questions raised about the general failure to report other variables in the student samples, such as intelligence and the socioeconomic status of the schools where samples were taken. It is rare that intelligence is even mentioned in this book and the schools used are vaguely described as in or near New Haven. In one experiment, out of three classes, one was suspected to be of lower intelligence, but this was dismissed as unimportant because the statistics would be of the second order. Nevertheless, when the over-all achievement of the three classes indicated that the suspected class indeed did not do as well as the other two, this was pointed out. In another experiment the classes were carefully matched as to intelligence and other factors. These few references stand out against most of the reports where no mention is made of these factors. One could wish also that there was complete uniformity in reporting significance of results explicitly by levels of confidence in order to make perfect the statistical picture.

These faults cannot detract from the impressiveness of the Yale Motion Picture Project. Any serious student of educational films will need to study these experiments in detail. The methodology is particularly important, representing the kind of design and statistical analysis which should be required in all film research today. These experiments bring us closer to a real insight into what makes films effective. Together with the comparable studies

at Pennsylvania State University, they have raised the level of experimentation in the audio-visual field to command attention and respect. Only such large-scale coordinated research programs seem to be able to command the financing, staff, and facilities to provide a continuity of attack right through to publication. It is hoped that the National Defense Education Act of 1958 under Title VII will provide more of such meaningful research.



RESEARCH ABSTRACTS

TELEVISION

MYERS, LAWRENCE, JR. *The Impact of an Educational Television Program on Its Audience—"Books and Ideas."* Radio and Television Center, Syracuse University, Syracuse, New York. February 1959. 156 p.

Purpose: To assess the impact of an educational program series titled "Books and Ideas" by examining the relative effectiveness of a number of promotional techniques, mail as an indicator of audience characteristics, size of audience in relation to competing programming, cumulative viewing patterns, description of viewers, and estimates of program effectiveness in terms of a number of variables.

Procedure: This study incorporated audience building, audience size measurement, and audience analysis. Publicity for the series was extensive. The station gave spot announcements to call attention to the program. Libraries and clubs were involved. The series was broadcast on 20 consecutive Sundays. A telephone survey was made of listeners. Samples of viewers and nonviewers were given a 10-page personal interview questionnaire. A special three-page questionnaire was given to viewers to obtain information on the program.

Results: All audience-building efforts were of value, but many discovered the program by chance. Audience size was influenced more by the hour, weather, competitive programming, and preceding programs than by program content. The average viewer saw 7 of the 20 programs. It was not a small loyal audience but a large variable audience. Viewers differed significantly from nonviewers. 70 percent of viewers had attended college. The telephone panel method was found to be efficient and valid.—*L. Twyford*

FILM

WHITMYRE, JOHN W. "Psychiatric Patients' Audience Reactions to Types of Motion Pictures." *Journal of Clinical Psychology* 14: 259-64; July 1958.

Purpose: To determine whether there is any discernible relationship between the type of movie shown a psychiatric patient audience and the degree of overt disturbance which the audience shows while viewing the picture.

Procedure: Two of each of the following types of films were shown to two groups of VA Hospital psychiatric patients: Western, adventure, comedy,

drama comedy, and war movies. (War movies is one category of several banned for use in VA hospitals, but an exception was made for the purposes of this experiment.) The two groups of VA patients were those from an open ward who had good contact with reality and those from a closed ward who had poor contact with reality. A 14-point rating scale was constructed by having two independent sets of judges rank the 14 kinds of overt behavior, such as talking or sleeping, which could be observed during the showing of the films. The ranks were converted to scale values by the method of Guilford. While the films were being shown, two observers made a record once each minute of the type of reaction of each of 12 subjects. The distribution of the average disturbance scores of the patients was found to be markedly skewed, because the patients had zero scores for some movies, i.e., showed no instance of disturbed behavior. The scores were, therefore, converted to *T*-scores before being treated by the analysis of variance in the method of Edwards. Analyses were made of the variation between wards, between subjects in the same ward, between movies, the interaction of movies and wards, and the interaction pooled *S*'s and trials.

Results: Only one *F*-test was significant at the 5 percent level; this was the one between wards. The patients from the open ward, those in better contact with reality, showed more evidence of disturbance. The author points out that this was not unexpected, since these patients would be more responsive to the environment. This difference between types of patients did not reappear when the experiment was repeated with another series of similar films. The author concludes that movies do not differ systematically in their tendency to elicit disturbances in the patients in general, nor do certain kinds of movies disturb certain kinds of patients. He questions, therefore, the necessity of banning films to be shown to patients. His results are supported by nurses' reports that on the nights following the showing of the banned war movies there was no increase in patient restlessness in the wards in question.—*Paul Wendt*

PHOTOGRAPHS

PETTIGREW, THOMAS F.; ALLPORT, GORDON W.; and BARNETT, ERIC O. "Binocular Resolution and Perception of Race in South Africa." *British Journal of Psychology* 49: 265-78; November 1958.

Purpose: To determine if members of different ethnic groups will vary in their racial identification of each other.

Procedure: Stimulus material consisted of similarly posed, full-faced, bust photographs of both sexes and of all four major ethnic groups of South Africa: Africans, Coloreds, Indians, European whites. 40 such pictures were used: 20 of each sex and 10 of each race. They were shown in pairs through a binocular arrangement so the subject saw the left picture with his left eye and the right picture with his right eye. In the practice series, identical pictures were shown to the two eyes. In the experimental series, pairs were formed of different races and different sexes. The pairs were equalized as to pose, age, sex, and pleasantness of expression, in order to maximize the likelihood of the

fusion of the binocular stimuli—that it, so the subject would tend to see a single fused picture instead of two separate ones. The exposure was limited to two seconds. Immediately thereafter, the subject was asked to name the race he thought he saw represented. At the end of the series it was repeated with the two photographs reversed from left to right in order to offset the effect of eye dominance. 122 subjects representing five different ethnic groups participated in the study. These were residents of Durban and vicinity. The authors admit that the South African social setting made it impossible to obtain a homogeneous social-economic group of subjects, but feel they were representative.

Results: All subjects reached a decision in less than two seconds with a high subjective certainty concerning which one of the four races was represented by the binocular stimuli. There was no evidence of conflict; a single judgment was given promptly. When different races were presented to the two eyes they were seen most often as "intermediate"—Coloreds or Indians. But the subjective determinants caused observers to be most accurate in judging members of their own group. The Afrikaners, who constituted part of the European group, evidenced a "suppression" response, in that they tended to identify the mixed pairs of photographs either as European whites or Africans. The authors believe that more study should be made involving binocular vision and the tachistoscopic presentation.—*Paul Wendt*

TEACHING MACHINES

TWYFORD, L. C.; O'HARA, J. G.; and GOLDMAN, A. *Green Light Rater.* Technical Report NAVTRADEVCEN 20-OS-3-1. U. S. Naval Training Devices Center, Port Washington, New York. December 23, 1958. 5 p.

Purpose: The purpose of this research and development project was to design and evaluate a teach-test, or self-rating, device which would incorporate good psycho-educational principles and which would permit the user to prepare and change the questions.

Procedure: A device was designed to provide teaching and practice to supplement more formal classroom instruction. It was designed to be inexpensive, adaptable to the physical environment of the trainee, and require minimal maintenance. Psychological factors involved in its design included (a) incentive motivation, (b) knowledge of results—immediate, specific, and differential, (c) repetition, (d) contiguity, (e) distribution of trials, (f) incidental learning, and (g) lack of stress. The apparatus displays 42 cards containing multiple-choice questions. One of four buttons beneath each card glows green if correctly pressed; the others glow red to indicate errors. Field evaluators answered a number of questions regarding the degree of use, maintenance problems, reactions of trainees, et cetera.

Results: This rater was found to have the advantage of similar teach-test devices. Production costs were estimated to be below those of similar devices; it was easy to operate and maintain. With two modifications it was recommended for production.—*L. Twyford*

The Library of Congress has cataloged this publication as follows:

Audio visual communication review. v. 1–
winter 1953–
[Washington]

v. in illus. 23 cm. quarterly.

Published by the Dept. of Audio-Visual Instruction, National Education Association of the United States.

1. Communication—Period. I. National Education Association of the United States. Dept. of Audio-Visual Instruction.

P90.A85 371.33305 56–4703

Library of Congress [5]

SPECIAL ISSUE OF AUDIO-VISUAL COMMUNICATION REVIEW

GRAPHIC COMMUNICATION AND THE CRISIS IN EDUCATION. Paper bound $1.25; cloth bound $2.00. Discounts on quantity orders apply.

OTHER PUBLICATIONS

of the

Department of Audio-Visual Instruction
National Education Association of the United States
1201 Sixteenth Street, Northwest, Washington 6, D. C.

A Directory of Graduate AV Programs for the Professional Education of Audio-Visual Supervisors, Directors, and Building Coordinators, $0.50

National Tape Recording Catalog, Second Edition, $1.00; 1958 Supplement, $0.50

The School Administrator and His Audio-Visual Program—1954 Yearbook, $3.75

Television in Instruction: An Appraisal, $1.00

The Cooperative Approach to Audio-Visual Programs, $1.50

Planning Schools for Use of Audio-Visual Materials—No. 1, Classrooms, Third Edition, $1.50

Planning Schools for Use of Audio-Visual Materials—No. 2, Auditoriums, $1.00

Planning Schools for Use of Audio-Visual Materials—No. 3, The AV Instructional Materials Center, $1.00

Planning Schools for Use of Audio-Visual Materials—No. 4, Audio-Visual Centers in Colleges and Universities, $1.50

Proceedings of DAVI Television Seminar, $1.00

Membership in
Department of Audio-Visual Instruction
Including subscription to AUDIOVISUAL INSTRUCTION
$5.00

All publications, except Audio-Visual COMMUNICATION REVIEW, are subject to the following discounts on quantity order:

2 to 9 copies	10%
10 or more copies	20%
For resale purposes	20%

Orders for $1.00 or less must be accompanied by a check or money order.